

From: Swope, Sheridan
Sent: Thursday, June 09, 2005 6:04 PM
To: STIC-Biotech/ChemLib
Subject: 10023515

For 10023515, pls search and interference search:

SID 1 against the NT and AA data bases.

SID 2 against the NT and AA data bases.

SID 3 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

Toby
Port

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 08:57:45 / Search time 89 Seconds
(without alignments)
3342.899 Million cell updates/sec

Title: US-10-023-515-2
Perfect score: 3079
Sequence: 1 MPQGLTSSASQWCFLLIQP.....PLSLTFLSLIQPFPPFCAP 581

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_tramb1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2915	94.7	575 2	Q6NTJ32
2	2606	84.6	525 2	Q6NDN9
3	2454	79.7	469 2	Q8NDC8
4	2250	73.1	575 2	Q6AM47
5	2094.5	68.0	575 2	Q6AM46
6	1985	64.5	545 2	Q81034
7	1676.5	54.4	361 2	Q95KH3
8	1298.5	42.2	557 1	SASB_ANAPL
9	1282.5	41.7	561 2	Q91WGO
10	1281	41.6	559 2	Q8BK48
11	1276.5	41.5	559 2	Q35533
12	1271	41.3	559 2	Q8BM97
13	1266.5	41.1	556 2	Q6PDB7
14	1263.5	41.0	568 2	Q6GMS4
15	1262.5	41.0	561 2	Q70177
16	1260.5	40.9	560 2	Q8R097
17	1257.5	40.8	554 2	Q35535
18	1239.5	40.3	562 2	Q6P306
19	1238	40.2	558 2	Q8K3R0
20	1233.5	40.0	558 2	Q8Q2R3
21	1231.5	40.0	532 1	ESR2_RABIT
22	1230	39.9	559 1	ESR2_HUMAN
23	1213.5	39.4	561 2	Q70631
24	1213.5	39.4	566 2	Q766D7
25	1213.5	39.4	566 2	Q864S9
26	1210.5	39.3	557 2	Q640T6
27	1199	38.9	561 1	ESR1_MESAU
28	1191	38.7	586 2	Q61PK9
29	1187.5	38.6	568 2	Q6PG74
30	1187.5	38.6	572 1	Q810S9
31	1186.5	38.5	554 1	ESTM_MOUSE

32	1183	38.4	565 2	Q95N05	Q95N05 canis fam1
33	1179.5	38.3	571 2	Q6UW8	Q6UW8 homo sapien
34	1179	38.3	568 2	Q726J1	Q726J1 homo sapien
35	1175	38.2	565 2	Q8RD29	Q8RD29 homo sapien
36	1171.5	38.0	568 2	Q8VCU1	Q8VCU1 mus musculu
37	1169.5	38.0	561 2	Q91WU0	Q91WU0 mus musculu
38	1164	37.8	565 2	Q97582	Q97582 sus scrofa
39	1162.5	37.8	555 2	Q6GML1	Q6GML1 brachydanto
40	1162.5	37.8	566 1	ESR1_PIG	Q29550 sus scrofa
41	1159.5	37.7	566 2	Q464Z1	Q464Z1 macaca fasc
42	1150.5	37.4	565 1	ESR1_RABIT	P12337 oryctolagus
43	1150	37.3	554 1	ESTN_MOUSE	P23953 mus musculu
44	1149	37.3	565 2	Q35534	Q35534 mesocricetu
45	1144	37.2	549 1	ESR1_RAT	P10959 rattus norv

ALIGNMENTS

RESULT 1

ID Q6NTJ32 PRELIMINARY; PRT: 575 AA.

AC Q6NTJ32; 05-JUL-2004 (TRENBLREL. 27, Created)

DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)

DE FLJ131547 protein.

GN Name=FLJ131547.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RA [1]

RP SEQUENCE FROM N.A.

RC TISSUE=PCR rescued clones;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loguettano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huix S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RM [2]

RP SEQUENCE FROM N.A.

RC TISSUE=PCR rescued clones;

RX Director MGC Project;

RA Submitted (Apr-2004) to the EMBL/Genbank/DBJ databases.

RT -1 SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.

EMBL: BC069501; AAH69501.1; -.

DR HSSP; P12337; IKAY.

DR GO; GO:0003824; F:catalytic activity; IEA.

DR InterPro; IPR002018; Carboxylesterase.

DR Pfam; PF00135; Coesterase; 1.

DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.

DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.

KW Hydrolase.

SQ SEQUENCE 575 AA; 63926 MW; 6F5B735BDFC9C09 CRC64;

Query Match 94.7% Score 2915; DB 2; Length 575;
 Best Local Similarity 97.0%; Pred. No. 6.2e-214;
 Matches 553; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 12 WCFLLIQPLIGHRQWKGKTPSGAAGPQNRRLGNIQKQVTVLGSPPVAVVFLGVPFPAAP 71
 DB 14 MAIWLVLAP-----TKGPSAGPQRNRLGNIQKQVTVLGSPPVAVVFLGVPFPAAP 65
 QY 72 PLGSLRPTNPQAPSPMONTREATSYRPLCTQNSEMLLDQMLKVHPKFGVSEDCLYIN 131
 DB 66 PLGSLRPTNPQAPSPMONTREATSYRPLCTQNSEMLLDQMLKVHPKFGVSEDCLYIN 125
 QY 132 IYAPAHADTGSKLPLVWPFPGAFKTSASIFDGSALAAVEDLVVVVQVRLGIFGFFTT 191
 DB 126 IYAPAHADTGSKLPLVWPFPGAFKTSASIFDGSALAAVEDLVVVVQVRLGIFGFFTT 185
 QY 192 WDQHAAGNNAFKQVVAALSVQKNIEPFGDPSSVTIIFGSSAGASVSSLLISPMKGLF 251
 DB 186 WDQHAAGNNAFKQVVAALSVQKNIEPFGDPSSVTIIFGSSAGASVSSLLISPMKGLF 245
 QY 252 HKAIMSGVAIIIPYLAHBYEKSEDLQVVAHFGGNNASSEALRLCTRKPSKELLTSLQ 311
 DB 246 HKAIMSGVAIIIPYLAHBYEKSEDLQVVAHFGGNNASSEALRLCTRKPSKELLTSLQ 305
 QY 312 KTKSFTFRVVDGAFEPNEPLDLISQKAFKAIPLIIGVNNHECGFLPMKEAPFLLSGSNKS 371
 DB 306 KTKSFTFRVVDGAFEPNEPLDLISQKAFKAIPLIIGVNNHECGFLPMKEAPFLLSGSNKS 365
 QY 372 LALHLIQNLHLIPROVLHIVANEYFPHDKSLTEIRDSLLDLGDVFFVVPALITARYRD 431
 DB 366 LALHLIQNLHLIPROVLHIVANEYFPHDKSLTEIRDSLLDLGDVFFVVPALITARYRD 425
 QY 432 AGAPVYFEFRHRPOCFEDTKPAFYKADHADVRFVFGGAFKGLDIWPEGAETEEKLIS 491
 DB 426 AGAPVYFEFRHRPOCFEDTKPAFYKADHADVRFVFGGAFKGLDIWPEGAETEEKLIS 485
 QY 492 RRMCMYATPATGPNNGNDLSLMPAYNLTEQYLOLDINMSIGQLKSPRYDFTWTSTPL 551
 DB 486 RRMCMYATPATGPNNGNDLSLMPAYNLTEQYLOLDINMSIGQLKSPRYDFTWTSTPL 545
 QY 552 ILSASDMLHSPSLTFLSLQPPFFFCAP 581
 DB 546 ILSASDMLHSPSLTFLSLQPPFFFCAP 575

RESULT 2
 Q96DN9 PRELIMINARY; PRT; 525 AA.
 AC Q96DN9;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE Hypothetical protein FLJ31547.
 GN Name=FLJ31547;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eumetazoa; Primates; Catarrhini; Hominiidae; Homo.
 NCBI_TaxID=9606;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Ohtsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita M.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosioki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimura M., Watanabe M., Hirokawa S., Chiba Y., Ishida S.,
 RA Oono Y., Takiguchi S., Watanabe S., Yosida M., Houchi T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,

RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musahino K., Yuki H., Oshima A., Sasaki K., Arita M., Sano S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe K., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Oono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura Y., Kikuchi H., Maeno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Igogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.",
 RL Nat. Genet. 36:40-45 (2004).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE-PCR rescued clones;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stelpien M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M.J., Udell T.B., Toshyunk J., Abramson R.D., Mullany S.J.,
 RA Raha S.S., Loggiano N.A., Peters G.V., Carminci P., Prange C.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Keteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting R.M., Madan A., Young A.C., Green E.D., Dickson M.C.,
 RA Blakeley R.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Rodriguez A.C., Skalka U., Smalins D.E., Scherch A., Schein J.E.,
 RA Krzyzanski M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.O., Warr M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PCR rescued clones;
 RA Strausberg R.,
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- Similarity: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL: AK056109; BAB71094.1; -
 DR EMBL: BC069548; BAB69548.1; -
 DR HSSP: P12337; 1k4Y.
 DR GO: GO:0016787; F:hydrolyase activity; IEA.
 DR InterPro: IPR002018; Carboxylesterase.
 DR InterPro: IPR000379; Ser. esters.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE: PS00122; CARBOXYL ESTERASE B_1; 1.
 DR PROSITE: PS00941; CARBOXYL ESTERASE B_2; 1.
 KW Hydrolyase, Hypothetical protein.
 SQ SEQUENCE 525 AA; 58200 MW; 7724878B8A22F215 CRC64;

Query Match 84.6% Score 2606; DB 2; Length 525;
 Best Local Similarity 90.9%; Pred. No. 2.2e-190;
 Matches 501; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

QY 31 GPSAGPQRNRLGNIQKQVTVLGSPPVAVVFLGVPFPAAPLGSRLFTNPQAPSPMDNL 90
 DB 25 GPSAGPQRNRLGNIQKQVTVLGSPPVAVVFLGVPFPAAPLGSRLFTNPQAPSPMDNL 84
 QY 91 REATSPMLCTQNSEMLLDQMLKVHPKFGVSEDCLYINTIYAPAHADTGSKLPLVWPF 150
 DB 85 REATSPMLCTQNSEMLLDQMLKVHPKFGVSEDCLYINTIYAPAHADTGSKLPLVWPF 144

```

QY 151 PGGAFTKGSASIFDGSALAAVEDLVVVVVOYRLGIFGFTTMDQHPGNMAFKQVAAVS 210
DB 145 PGGAFTKGSASIFDGSALAAVEDLVVVVVOYRLGIFGFTTMDQHPGNMAFKQVAAVS 204
QY 211 WVQKRIEFGGDPSSVTTFGSGAGAISVSLILSPMAKGLFHKAIMEGVAIIPYLEAHD 270
DB 205 WVQKRIEFGGDPSSVTTFGSGAGAISVSLILSPMAKGLFHKAIMEGVAIIPYLEAHD 264
QY 271 YEKESDQVVAHPGCGNNADEBALRCRTKPSKELLTSLQKTSFPTVVDGAFPEPNEPL 330
DB 265 YEKESDQVVAHPGCGNNADEBALRCRTKPSKELLTSLQKTSFPTVVDGAFPEPNEPL 324
QY 331 DLSQKAFKAIPIIIGVNNHCGFLIPMKKAPILSGSNKSLALHLIQLNIHPIPOYLHL 390
DB 325 DLSQKAFKAIPIIIGVNNHCGFLIPMKKAPILSGSNKSLALHLIQLNIHPIPOYLHL 384
QY 391 VANSEYFHDKSLTEIRDSLLDLGDPFVVPALITARYHDAAGAVYIEFRHRPOCED 450
DB 385 VANSEYFHDKSLTEIRDSLLDLGDPFVVPALITARYHDAAGAVYIEFRHRPOCED 424
QY 451 TKPAPVADHADVRFVFGGAFPKGDIMFEGATEEKEELSRKMKKWTATPRTGNPNCN 510
DB 425 -----EGATEEKEELSRKMKKWTATPRTGNPNCN 454
QY 511 DLSMPAYNLTEOYLQDLNMSLQRLKPRVDFTSTIPILSASDMLHSPILSLTFLS 570
DB 455 DLSMPAYNLTEOYLQDLNMSLQRLKPRVDFTSTIPILSASDMLHSPILSLTFLS 514
QY 571 LLOPFFFCAP 581
DB 515 LLOPFFFCAP 525

```

RESULT 3

```

Q8NBC8 PRELIMINARY; PRT; 469 AA.
AC Q8NBC8;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ33678.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai T., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E., Omura Y.,
RA Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Houta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotake S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Teraishi Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki M., Watanabe K., Kumagai A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,

```

```

RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Mizumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: AK090997; BAC03565.1; -
DR HSP: P12337; 1K4Y.
DR GO: GO:0016787; F:Hydrolase activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR00379; Ser_esterase.
DR Pfam: PF00135; Coesterase_1.
DR PROSITE: PS00122; CARBOXYLESTERASE B 1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
DR KEGG: K00001; CARBOXYLESTERASE_B_2; 1.
SQ SEQUENCE 469 AA; 52312 MW; 7A18BF68C10A9080 CRC64;

```

Query Match 79.7%; Score 2454; DB 2; Length 469;

Best Local Similarity 99.6%; Pred. No. 7, 8e-179; Mismatches 0; Indels 0; Gaps 0;

Matches 467; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

QY 113 MLKHYPRFVSEDCLYINITYAPAHADGSKLPVLWPFPGAFTKGSASIFDGSALAAVE 172
DB 1 MLKHYPRFVSEDCLYINITYAPAHADGSKLPVLWPFPGAFTKGSASIFDGSALAAVE 60
QY 173 DVLVVVVOYRLGIFGFTTMDQHPGNMAFKQVAAVSWQKRIEFGGDPSSVTTFGSG 232
DB 61 DVLVVVVOYRLGIFGFTTMDQHPGNMAFKQVAAVSWQKRIEFGGDPSSVTTFGSG 120
QY 223 AGAISVSSLLSPMAKGLFHKAIMEGVAIIPYLEAHDYKESDQVVAHPGCGNNADE 292
DB 121 AGAISVSSLLSPMAKGLFHKAIMEGVAIIPYLEAHDYKESDQVVAHPGCGNNADE 180
QY 293 ALLRCLRTKPSKELLTSLQKTSFTRVVDGAFPEPNEPLDLSQKAFKAIPIIIGVNNH 352
DB 181 ALLRCLRTKPSKELLTSLQKTSFTRVVDGAFPEPNEPLDLSQKAFKAIPIIIGVNNH 240
QY 353 GFLPMKAPRILSGSNKSLALHLIQLNIHPIPOYLHVANSEYFHDKSLTEIRDSLL 412
DB 241 GFLPMKAPRILSGSNKSLALHLIQLNIHPIPOYLHVANSEYFHDKSLTEIRDSLL 300
QY 413 LGDVFVVPALITARYHDAAGAVYIEFRHRPOCEDTKPAPVADHADVRFVFGGAF 472
DB 301 LGDVFVVPALITARYHDAAGAVYIEFRHRPOCEDTKPAPVADHADVRFVFGGAF 360
QY 473 LKGDIVMEGATEEKEELSRKMKKWTATPRTGNPNCNLSMPAYNLTEOYLQDLNMS 532
DB 361 LKGDIVMEGATEEKEELSRKMKKWTATPRTGNPNCNLSMPAYNLTEOYLQDLNMS 420
QY 533 LQRIKPRVDFTSTIPILSASDMLHSPILSLTFLSLOPFFFCAP 581
DB 421 LQRIKPRVDFTSTIPILSASDMLHSPILSLTFLSLOPFFFCAP 469

```

RESULT 4

```

Q6AM47 PRELIMINARY; PRT; 575 AA.
AC Q6AM47;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=cauxin;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxId=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Miyazaki M., Yamashita T., Taira H., Suzuki A.;
RT "cauxin family protein.";

```

RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: AB186392; BAD35015.1; -
DR GO: GO:0003824; F:catalytic activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR000379; Ser_estrs.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydroxylase.
SQ SEQUENCE 575 AA; 63620 MW; 1005C35B82B1183D CRC64;

Query Match 73.1%; Score 2250; DB 2; Length 575;
Best Local Similarity 75.7%; Pred. No. 4e-163;
Matches 417; Conservative 61; Mismatches 73; Indels 0; Gaps 0;

```
QY 31 GPSAEGPQNRTRLGWQKQVTVLGSPPVNVVFLGSPFAAPPLGSGIRFPNPPDASPMNDL 90
DB 25 GRPAADPVPVSTRIGWVGRQATVLTSTWVNVFLGIPFAAPPLGPRFRPRPALLMND 84
QY 91 REATSYPNLCLQNSEWLLDQHLKVHYPKFVSEDCLYLNTYAPAHADTSGKLPVLYWF 150
DB 85 RDATSYPRKCLQNSWLLSDQHLKVHYNLEVEDCLYNTYAPAHANTSGKLPVLYWF 144
QY 151 PGCAFKTGSASTFDGSALAAVEDVLYVTVQRLGIFGFFTTWDQAPGNMAFKDQVAALS 210
DB 145 PGCAFKTGSASTFDGSALAAVEDVLYVTVQRLGIFGFFTTWDQAPGNMAFKDQVAALS 204
QY 211 WQKNIEFFGDPSSVTIFGSGAGASVSSLILSPMAKGLFHKALIMESGVAILPYLEAHD 270
DB 205 WQKNIEFFGDPSSVTIFGSGAGASVSSLILSPMAKGLFHKALIMESGVAILPYLEAHD 264
QY 271 YEKSDLVQVAHFCGNNAASDEBALLRCLRTKPSKELLTLOKTKSFTRVVDGAFEPNEPL 330
DB 265 DERNDLVQVARIICGNVSDSVALLQCLRAKSSSELDLIDNKTKSFTRVVDGAFEPNEPL 324
QY 331 DLLSKAKAIPISTIGVNNHCEGFLPMKEAPELISGNSKSLAHLNIIHIPQYLAH 390
DB 335 DLLSEKTNISIPSVIGVNNHCEGFLPMKEAPELISGNSKSLAHLNIIHIPQYLAH 384
QY 391 VANEFYHDKSLTEIRDSLLDLGDVFFVVPALITARYHDAAGAVVYEFERHPOCED 450
DB 385 VADQYFYHDKSPVEIRDSFLDLGDVFFVVPALITARYHDAAGAVVYEFERHPOCED 444
QY 451 TKPAFVADHADVRFVFGAFLKGDIVMEFGATEBEKLSRRKMKWTATPARTGNNGN 510
DB 445 TRPAFVADHADVRFVFGAFLKGNIVMEFGATEBEKLSRRKMKWTATPARTGNNGN 504
QY 511 DLSTMPAYNLTEQYLOLDLNMISLGORLKEPRVDFTSTPIPLISASDMLSPSLSTFLS 570
DB 505 GLPMPAYNSQEQYLIKLDLNMISVGOKLKEQVEFWSDTLPILMSMSTAPGPPVPLSLIS 564
QY 571 LLQPPFFFCAP 581
DB 565 VLPFLFESSAP 575
```

RESULT 5

Q6AM46 PRELIMINARY; PRT; 575 AA.
AC Q6AM46;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=cauxin;
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN RP SEQUENCE FROM N.A.
RA Miyazaki M., Yamaehita T., Taira H., Suzuki A.;

RT "cauxin family protein.";
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: AB186393; BAD35016.1; -
DR GO: GO:0003824; F:catalytic activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR000379; Ser_estrs.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydroxylase.
SQ SEQUENCE 575 AA; 64166 MW; 3EB9D85981D9DE0A CRC64;

Query Match 68.0%; Score 2094.5; DB 2; Length 575;
Best Local Similarity 68.5%; Pred. No. 3e-151;
Matches 390; Conservative 69; Mismatches 99; Indels 11; Gaps 1;

```
QY 12 WCFPLIOPILGHRQWKTGPSAQPNRTRLGWQKQVTVLGSPPVNVVFLGIPFAAP 71
DB 17 WIFGALIE-----GSYIEPRHRYTKLGWQKQATVLTGRLEPVNVFLGIPFAAP 65
QY 72 PLGSLRTNPPDASPMNDLEATSYPNLCLQNSEWLLDQHLKVHYPKFVSEDCLYLN 131
DB 66 PLGSLRFSKQPPPIPMNDLREATAVYPLCFQNLLEWLFYQNLTKVSPILGMSDCLYLN 125
QY 132 IYAPAHADTSGKLPVLYWFPGCAFKTGSASTFDGSALAAVEDVLYVTVQRLGIFGFFTT 191
DB 126 IYAPAHADTSGKLPVLYWFPGCAFKTGSASTFDGSALAAVEDVLYVTVQRLGIFGFFTT 185
QY 192 WDQAPGNMAFKDQVAALSWQKNIEFFGDPSSVTIFGSGAGASVSSLILSPMAKGLF 251
DB 186 ONQHAFGNMAFWQDLAALLVRENIRYFGGNPSPVITIFGNSAGATISLILSPSADLF 245
QY 252 HKAIMESGVAILIYLAHDEKSEDIQVAHFCGNNAASDEBALLRCLRTKPSKELLTSLQ 311
DB 246 HRAIMESGVAILIYLAHDEKSEDIQVAHFCGNNAASDEBALLRCLRTKPSKELLTSLQ 305
QY 312 KTSFTRVVDGAFEPNEPLDLLSKAKAIPISTIGVNNHCEGFLPMKEAPELISGNS 371
DB 306 KAKSFTRVVDGAFEPNEPLDLLSKAKAIPISTIGVNNHCEGFLPMKEAPELISGNS 365
QY 372 LAHLNIIHIPQYLAHVANEFYHDKSLTEIRDSLLDLGDVFFVVPALITARYHD 431
DB 366 TALTLIHTLHIPTQLYIVTKYEFHGKSPDIDRTLLDLGDVFFVVPALITARYHD 425
QY 432 AGAPVYFYEFRHPOCEDTKPAFYKADHADVRFVFGAFLKGDIVMEFGATEBEKLS 491
DB 426 SGAPVYFYEFRHPOCEDTKPAFYKADHADVRFVFGAFLKGDIVMEFGATEBEKLS 485
QY 492 RKMKTWATFARTGNNGNDLSLMPAYNLTEQYLOLDLNMISLGORLKEPRVDFTSTPI 551
DB 486 RKMKTWATFARTGNNGNDLSLMPAYNLTEQYLOLDLNMISLGORLKEPRVDFTSTPI 545
QY 552 ILSASDMLSPSLSTFLSLLQPPFFFCAP 580
DB 546 ILSASDMLSPSLSTFLSLLQPPFFFCAP 574
```

RESULT 6

O81034 PRELIMINARY; PRT; 545 AA.
AC O81034;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=cauxin;
OS Fells silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedta; Felidae; Fells.
OX NCBI_TaxID=9685;
RN RP SEQUENCE FROM N.A.

RC TISSUB=Kidney; PubMed=12401131; DOI=10.1042/BJ20021446;
 RA MEDLINE=22458314; Kamie K., Soeta S., Taira H., Yamashita T.;
 RT "Molecular cloning and characterization of a novel carboxylesterase-
 RT like protein that is physiologically present at high concentrations in
 RT the urine of domestic cats (Felis catus).";
 RL Biochem. J. 370:101-110(2003).
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AB045377; BAC22577.1; -.
 DR HSSP; P12337; 1K4Y.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR InterPro; IPR002018; Carboxylesterase.
 DR InterPro; IPR000379; Ser. estera.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE B 1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 DR Hydrolyase.
 KW SEQUENCE. 545 AA; 60505 MW; 9F73FA693D271FA9 CRC64;
 SQ
 Query Match 64.5%; Score 1985; DB 2; Length 545;
 Best Local Similarity 71.7%; Pred. No. 6,5e-143;
 Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;
 QY 31 GPSAEGPQNRTRLGWIQGVYVIGSPVPVNVFLGVPAPPLGSLRFTNPQAPSPMDNL 90
 DB 25 GPADAPVRSYRLGVNRGKQTTVLGASTVPVNMFLGIPYAAPPLGLRFPKQKPPALPGNDP 84
 QY 91 REATSPYRLCQNSEWMLLDQHLKXHPKFGVSDCLYNITYAPAHADTSGKLPVLYWF 150
 DB 85 RNATSTYPLCLCPQDLEMVSYQHVLRVYPKLEASDCILYNIYAPAHADNCSNLPVWVWF 144
 QY 151 PGAFKTKGSASIPGSAALAYEDVLYVVOYRLGIFGFTTMDQAPGNMAFKQVALS 210
 DB 145 PGAFKTKGSASIPGSAALAYEDVLYVVOYRLGIFGFTTMDQAPGNMAFKQVALS 204
 QY 211 WVQKNIEPFGDPSVITIFGSSAGASIVSSSLISPMAGLFRKAIMESGVAIIPYL-BAH 269
 DB 205 WVRNIEPFGDPSVITIFGSSAGASIVSSSLISPMAGLFRKAIMESGVAIIPYL-BAH 264
 QY 270 DYKSEBDQVVAHFCGNNASDSEALRLCTKPKSEKLLTSQTKSPFRVVDGAFPNR 329
 DB 265 GDERKQDQVVARICGCHASDAPALLOCTLRAPKPSBELMDISKULTFSPVDDPFPDEP 324
 QY 330 LDLSOKAFKAIPTIIGVNHCEGFLPMKEAPBELSGSNSLALHLIQLNLIHPPOYLH 389
 DB 325 VALTIQKAFNSVPSIIIGVNHCEGFL-STEPSEILGSSNSLALYLVHTLNLNITPQYLH 383
 QY 390 LVANEYFPHDKSLTEIRDSLDDLGDVFPVVPALITATYHRDAGAPVYEFRRHQCFE 449
 DB 384 LVADHYFYNKSPVEIRDSFIDLGDVFPVVPVLTATYHRDAGAPVYEFRRHQCFE 443
 QY 450 DTQAPVAKADADDEVRFVFGGAPLKGDIWMFEGATEEKKLSRKMKTWATPARTGNDG 509
 DB 444 DTRAPVAVADSDERFVFGGAPLKGDIWMFEGATEEKKLSRKMKTWATPARTGNDG 503
 QY 510 NDLSLMPAYNLTBOYLQDLNMSLQGRKXPRVPTSTI 549
 DB 504 EGVPLMPAYTQSEYLLKDLISVSGQKKEQVEPMTNTI 543
 RESULT 7
 ID Q95KH3 PRELIMINARY; PRT; 361 AA.
 AC Q95KH3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopitheciinae; Macaca.
 NCBI_TaxID=9541;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Temporal lobe right;
 RA Oosada N., Hida M., Kuenda J., Tanuma R., Ieki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AB060873; BAB46884.1; -.
 DR HSSP; P12337; 1K4Y.
 DR InterPro; IPR002018; Carboxylesterase.
 DR Pfam; PF00135; Coesterase; 1.
 DR Hypothetical protein.
 KW SEQUENCE. 361 AA; 41032 MW; AC62AB6E78C5F1A1 CRC64;
 SQ
 Query Match 54.4%; Score 1676.5; DB 2; Length 361;
 Best Local Similarity 95.0%; Pred. No. 1.4e-119;
 Matches 323; Conservative 5; Mismatches 11; Indels 1; Gaps 1;
 QY 242 ILSPMAKGLFRKAIMESGVAIIPYLEADHYKESDQVVAHFCGNNASDSEALRLCTRTX 301
 DB 17 ILSPMAKGLFRKAIMESGVAIIPYLEADHYKESDQVVAHFCGNNASDSEALRLCTRTX 76
 QY 302 PSKELLTSQTKSPFRVVDGAFPNRBDLLSQAKFAITSIIGVNHCEGFLPMKEA 361
 DB 77 PSKELLTSQTKSPFRVVDGAFPNRBDLLSQAKFAITSIIGVNHCEGFLPM-BA 135
 QY 362 PEILSGNSKSLALHLIQLNLIHPPOYLHVNXYFHDGSLTEIRDSLDDLGDVFPVVP 421
 DB 136 PEILSGNSKSLALHLIQLNLIHPPOYLHVNXYFHDGSLTEIRDSLDDLGDVFPVVP 195
 QY 422 ALITARYHRDAGAPVYEFRRHQCFEDTTPAVKADHADDEVRFVFGGAPLKGDIWFE 481
 DB 196 ALITARYHRDAGAPVYEFRRHQCFEDTTPAVKADHADDEVRFVFGGAPLKGDIWFE 255
 QY 482 GATEEKKLSRKMKTWATPARTGNPNNGNDLSLMPAYNLTBOYLQDLNMSLQGRKLEPR 541
 DB 256 GATEEKKLSRKMKTWATPARTGNPNNGNDLSLMPAYNLTBOYLQDLNMSLQGRKLEPR 315
 QY 542 VDFWSTTPIILSASDMLHSPSLTFLSLQPFPPFCAP 581
 DB 316 VDFWSTTPIILSASDMLHSPSLTFLSLQPFPPFCAP 355
 RESULT 8
 SABS_ANAPL STANDARD; PRT; 557 AA.
 ID Q04791;
 AC Q04791;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Fatty acyl-CoA hydrolase precursor, medium chain (EC 3.1.2.-)
 DB (thioesterase B).
 OS Anas platyrhynchos (Domestic duck).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
 NCBI_TaxID=8839;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-65.
 RC TISSUE=Urological gland;
 RA MEDLINE=93300823; PubMed=8314791;
 RA Hwang C.-S., Kolatukudy P.E.;
 RT "Molecular cloning and sequencing of the thioesterase B cDNA and
 RT stimulation of expression of the thioesterase B gene associated with
 RT hormonal induction of peroxisome proliferation.";
 RL J. Biol. Chem. 268:14278-14284(1993).
 CC -1- FUNCTION: Fatty acid biosynthesis chain termination and release of
 CC the free fatty acid product is achieved by hydrolysis of the thio
 CC ester by a thioesterase. This thioesterase may be associated with
 CC peroxisome proliferation and may play a role in the production of
 CC 3-hydroxy fatty acid diester phenomones.
 CC -1- TISSUE SPECIFICITY: Highest levels in urological gland, much lower
 CC in liver and kidney.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L05493; AAA49223.1; -
 DR PIR: A47162; A47162.
 DR HSSP: O77540; 1K4Y.
 DR InterPro: IPR002018; CarboxylesteraseB.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
 KM Direct protein sequencing; Fatty acid biosynthesis; Hydrolase; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 557 Fatty acyl-CoA hydrolase.
 FT ACT_SITE 227 227 Acyl-ester intermediate (By similarity).
 FT ACT_SITE 345 345 Charge relay system (By similarity).
 FT ACT_SITE 460 460 Charge relay system (By similarity).
 FT DISULFID 93 122 By similarity.
 FT CARBOHYD 476 476 N-linked (GlcNAc...) (potential).
 SQ SEQUENCE 557 AA; 61637 MW; 03E35D9A037F6B0 CRC64;

Query Match 42.2%; Score 1298.5; DB 1; Length 557;
 Best Local Similarity 48.4%; Pred. No. 1.9e-90;
 Matches 265; Conservative 79; Mismatches 177; Indels 27; Gaps 8;

QY 30 TGPSAEGRQNRRLGIGQKQVTVGSPVNVFLGVPAPRPLGSLRFTNPQAPSPWDN 89
 DB 23 TGQRKEQEVVTVNYSGVQYKVAARSVVFLGFLPAKPVPLRFSBPQPEPWKG 82
 QY 90 LREATSYNLCIONSEWMLLDQHM-----KVHYKFCVSDSCLYNTIYAAHADTQSKL 144
 DB 83 VRDAASYPMCLQDK---VLQGYLSDATNRKRRKRLQISDCILNYTTFVSTEEDEKL 139
 QY 145 PVLVWFGAGFKTGSASIFDGSALAAVEDVLVVVQYRLGIFGFTTWDQAPGNMAFKD 204
 DB 140 PVFWVHGGGLVSGAASSYDGSALAFDNNVVVTIQYRLAGVFSQCDKARNMGLD 199
 QY 205 QVAALSWQKRIEFGGDSSTVTFGBSAGALSVSLSLSMAGLPHKAIMEGVAL-I 263
 DB 200 QVALLOMIQENTIHFRGDPGSVTTFGBSAGVSVALSLPLAKGLPHKAISEGTAVRI 259
 QY 264 PYLEAHDYKESDQVVAHFCGNNASDEALRLCIRTPSKEL--LTISQKTSFTVVD 321
 DB 260 LFTS-----QPEEQQRILAAAGCKSSALVVECLREKTEHMEQITLKNRPMFISLSD 315
 QY 322 GAFPNRPDLISQAKRAIISITGVNNHCEGFLP-MKEAPEILSGSNKSLAHLIQT 380
 DB 316 GVFEPPKSPRLSRKVINAVPYIIGVNNCEFGWILPMMKFEPTFEGLEKDVARQVLTST 375
 QY 381 LHI-----PROYLHLVANEYFHDGSLTETIRDSLLDGLGVFVPPVPLITARRHDA 435
 DB 376 LALSFKAPBDIVDLVNEIYIGVAENRAQVRDGLDSDIAPLFFVSIVVEARHRRDGNP 435
 QY 436 VYFYEPRHPOCFEDYKPAFVKADHAEVRFVFGAGLAKDIDVFEQATEEKLISKRM 495
 DB 436 VYFYEPRHPSAAGVVPFVKADHAEIAFVFGKPLAQN-----ATEEAKLRITW 489
 QY 496 KYWATFARNGPNPNDLSLWPAKYLTEQYQLDLNLSLGRLEKPRVDFTSTITPLILSA 555
 DB 490 KYTNFARNNGPNEGGLVHPQYDMERYLEIDLTKQAKKLEKRMKEFMWQLTEQIMSD 549
 QY 556 SDMLASPL 563
 DB 550 RRKHTDL 557

RESULT 9

Q91WGO ID Q91WGO PRELIMINARY; PRT; 561 AA.
 AC Q91WGO;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 26, Last annotation update)
 DE Carboxylesterase 2 (Carboxylesterase M3 precursor).
 GN Name=Ces2; Synonyms=ces2a3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon, and Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.J., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heide F.,
 RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Yoshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny K.C., Harte S., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywanski M.I., Skaleka U., Smalins D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RA Strausberg R.J.
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RA Strausberg R.J.
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon;
 RA Strausberg R.J.
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Colon, and Kidney;
 RA Strausberg R.J.
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RX MEDLINE=22744052; PubMed=12859986; DOI=10.1016/S0003-9861(03)00286-8;
 RA Furukawa T., Hosokawa M., Nakata F., Sacch T., Chiba K.,
 RT "Purification, molecular cloning, and functional expression of
 RT inducible mouse liver acylcarboxylate hydrolase in C57BL/6 mouse,
 RT belonging to the carboxylesterase multigene family.";
 RL Arch. Biochem. Biophys. 416:101-109 (2003).
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL: BC015290; AAH15290.1; -
 DR EMBL: BC024552; AAH24552.1; -
 DR EMBL: BC031170; AAH31170.1; -
 DR EMBL: BC034178; AAH34178.1; -
 DR EMBL: BC034180; AAH34180.1; -
 DR EMBL: BC034191; AAH34191.1; -
 DR EMBL: AB110073; BAC76623.1; -

Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywnicki M.I., Skalske U., Smalton D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

[6]
SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AK077248; BAC36707.1; -.
DR EMBL; BC055062; AAH55062.1; -.
DR HSSP; P12337; 1K4Y.
DR MGI; 2443170; 9030624L02Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR002018; Carboxylesterase.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR HydroLase.
SQ SEQUENCE 559 AA; 62317 MW; CC14C61034A122C3 CRC64;

Query Match 41.6%; Score 1281; DB 2; Length 559;
Best Local Similarity 48.7%; Pred. No. 4.2e-89;
Matches 260; Conservative 79; Mismatches 165; Indels 30; Gaps 8;

31 GPSAGPQRNTLGIQGVYVLSPPVNVVFLGVPAAPLGLRTNQPASPMNL 90
26 GQDASPIRNTHTGVRSLVAVKOTDIAVHTFLGIPFAKPPVGLRPAPEPMSGV 85
91 REATSYNLCIONSEWL-LDQMLKVHPKGVSEDCLYNITYAPAHADTGSKLPLYVW 149
86 RDGTSHPNMCIONDMLMGSEDLKMMNLILPISSEDCLYNITYAPAHAGSNLPLVW 145
150 FPGAFKTSASIPFGSALAAVEDLVVVOYRIGIFGFTTMDQAHAGNAFKDOVAL 209
146 IHGALTVMASMTDGSMLAATEDVVAIQIRLGLVGFSTGDOHAKGMNGYLDQVAL 205
210 SWVQKNIEFFGSDPSVTIFGESAGATSVSSLLISPMKGLFHKAIMSGVAILPYLEAH 269
206 RMVQONIVHFGNPDRTVITFESAGTSVSSHVSPMSGLFHKAIMSGVAVLPDLISS 265
270 DYKSEDLQVVAHFCGNNAASDEALLRLCTRPSEKELLTTSQKTSFTRVVDGAFPPNER 329
266 SSEVNH--RIYANISGCAAVNSEETLMCLRGKNEAEMLAINKVFIIPGVVDFLPKHP 323
330 LDLSQAKFAKIPSIIGVNNHCGFLP-----MKEAPEILSGNSKSLALH 376
324 QELMASDQFHVPSEIIGINDEYWIPLTMDPQKIEITRKLPVLV---KSTALKM 379
377 IONTLHLPQYLVHANEYFHDKSLTEIRDSLIDLIGVFFVVPALITARYRDGAPV 436
380 M-----LPREGDLMEEYMGDTEDPELLOAQFEMKGMDFMFIYPLQVANHFR-SHAPV 433
437 YFYEFRRHPCCFEDTKAFVKAHADAEVRFVFGGAFPKGDIWVPEGATEEKLSTRKMK 496
434 YFEYFQRHPEFDFRPVYKADHDEITFLVFGQF--GNIKL--PYTEEEEDLSRRIMK 489
497 YMATFARTGNPNNDLSLWPAVNLTEOYLQDLNMSLIGRLKEPRVDFTSTIP 550
490 YMANFARHGNPNEGFLPYFVPMVDHDEQYLQDLTOPSGRALKARRLQFTTKLP 543

RESULT 11
035533

ID 035533 PRELIMINARY; PRT; 559 AA.
AC 035533;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Carboxylesterase precursor (EC 3.1.1.1).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_Taxid=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Syrian golden; TISSUE=liver;
RA Some T., Wang C.Y.;
RT "Microsomal amidases and carboxylesterases.";
RL (in) Guengerich F.P. (eds.),
RL COMPREHENSIVE TOXICOLOGY VOLUME 3, Biochemical information, pp.265-281,
RL Pergamon, Oxford (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Syrian golden; TISSUE=liver;
RA Some T., Ishida Y., Takabatake E., Wang C., Isobe M.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; D50577; BAA23605.1; -.
DR HSSP; P12337; 1K4Y.
DR GO; GO:0004091; F:carboxylesterase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002018; Carboxylesterase.
DR InterPro; IPR000866; ER_target_S.
DR InterPro; IPR000379; Ser_estrs.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
DR HydroLase; signal.
FT SIGNAL 1 27 potential.
FT CHAIN 28 559 carboxylesterase.
SQ SEQUENCE 559 AA; 62126 MW; 5B6BD4F9A47C6AD8 CRC64;

Query Match 41.5%; Score 1276.5; DB 2; Length 559;
Best Local Similarity 49.6%; Pred. No. 9.2e-89;
Matches 262; Conservative 79; Mismatches 172; Indels 15; Gaps 9;

28 GKTPGABEPQNTLGIQGVYVLSPPVNVVFLGVPAAPLGLRTNQPASPM 87
26 GQDSE--PIRNTHTGVRSLVAVKOTDIAVHTFLGIPFAKPPVGLRPAPEPMP 82
88 DNLREATSYNLCIONSEWL-LDQMLKVHPKGVSEDCLYNITYAPAHADTGSKL 146
83 SGVRDGNSEFPMACIONDDMMNSBGLKMLTLPISSEDCLYNITYAPAHAGSNL 142
147 LVVFGGAFKTSASIPFGSALAAVEDLVVVOYRIGIFGFTTMDQAHAGNAFKDO 206
143 MWVINGALVVMGMAVMDGSMMLAIEDVVVVTIYRLGLVGFSTGDEHARGNNGYLD 202
207 AALSWQKNIEFFGSDPSVTIFGESAGATSVSSLLISPMKGLFHKAIMSGVAILPYL 266
203 AALRWQONIAHFGNPDQVITFESAGTSVSSHVSPMSGLFHKAIMSGVAVLPDLISS 262
267 EAHDEKESDLQVVAHFCGNNAASDEALLRLCTRPSEKELLTTSQKTSFTRVVDG 326
263 ISSSEEMV--YTIYANISDCAAVNTEVTLGLRKSSEALINKVFIIPGVVDFLP 320
327 NEPLDLSQAKFAKIPSIIGVNNHCGFLP--MKEAPEILSGNSKSL--ALHLIONLH 382
321 KHPDELMSADFPDVPSTIGVNNDEYGMVLVINGSAQEKIETRVVLPALIKSTANQ 380
363 IIPQYLVHANEYFHDKSLTEIRDSLIDLIGVFFVVPALITARYRDGAPVPEYFR 442
381 LPREGDLMEEYMGDNEDPELLOAQFEMKGMDFTFIYPLQVANHFR-AHAPYFTEFQ 439
443 HRPCCFEDTKAFVKAHADAEVRFVFGGAFPKGDIWVPEGATEEKLSTRKMKYMATFA 502


```

Db 440 HRPEFQKTRPHYADHGDSEFFPFGN-LLCG--IKF-ASTEEBELLSRQOMKMANFA 495
Oy 503 RTGNPNNDLSLMPAYNLTEOYLQDLNMSLQORLKEPVDFTSTIP 550
Db 496 RHGNPNSENLPYWPMLDHDQYLDIDIKPAVGRALKARRLOFTWTLIP 543

RESULT 12
Q8BM97 PRELIMINARY; PRT; 559 AA.
ID Q8BM97;
AC Q8BM97;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE Mus musculus adult male colon cDNA, RIKEN full-length enriched
DE library, clone:9030624102 product:similar to CARBOXYESTERASE (EC
DE 3.1.1.1) (ALI-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE)
DE (PROCAINE ESTERASE) (METHYLBUTYRASE).
GN Name:9030624102Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=99271253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA The PANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komano H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Komano H., Akizawa Y., Nishii K., Katsunaka T., Teshiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Washikaki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,

```

```

RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori P., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Komano H., Kouda M., Koya S.,
RA Kurihara C., Matsumoto T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akibira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yaeunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: AK033563; BAC28361.1; -.
DR HSSP: P12337; 1k4Y.
DR MGD: MGI:2443170; 9030624102Rik.
DR GO: GO:0005615; C:extracellular space; TAS.
DR InterPro: IPR002018; Carbesteraab.
DR InterPro: IPR000179; Ser_estre.
DR Pfam: PF00135; Coesterase_1.
DR PROSITE: PS00122; CARBOXYESTERASE B_1; 1.
DR PROSITE: PS00941; CARBOXYESTERASE_B_2; 1.
KM Hydrolyase.
SQ SEQUENCE 559 AA; 62305 MW; 8C54D21522C352C4 CRC64;

Query Match 41.3%; Score 1271; DB 2; Length 559;
Best Local Similarity 48.7%; Pred. No.2.4e-88;
Matches 260; Conservative 78; Mismatches 166; Indels 30; Gaps 8;

Oy 31 GPSAEGPQNRTRLMIGQKQVTVLGGPVPVNVFGLVPPAAPPLGSLRTNPQAPSWNL 90
Db 26 QGDSASPIRNTHTGQVRSILVHVVDIAVHTFGLIPAKKPPVGLRPAPDEAPPMGAV 85
Oy 91 REATSYPLCLQNSEWL-LDQHLKHYPRFGVSEDCLYNINYPARADGSLPVVW 149
Db 86 RDGTSHPMWCQDNLMKSEDLKMMNLILPISSEDCLYNINYPALAHGSLPVVW 145
Oy 150 FPGAFKTSASIFPGSALAYEDVLVVVQYRGIFGFTTDOHAGNNAFKQVVAL 209
Db 146 IHGHALTVGASMTDGSMLAATEDVYVAIQRLGVGFSTGQDHAGNMGYLDQVAL 205
Oy 210 SWVQNIIEFGGDPSSVTIFGESAGAISVSSLISPMKGLFHKAIMESGVAIIPLYEAH 269
Db 206 RMVQNIIVHFGANDRVITIFGESAGTSVSHVVSPMQGLFHGAIMESGVAIPLYLSS 265
Oy 270 DYKSESDIQVAHCGNNAASEALRLCLRTKPSKELLTSOKTSFRRYVDGAFFNBP 329
Db 266 SSEWVH--RIVANISGCAAVNSETLMLCCLRGKNEALEMIAINKVFRIIPGVVDGEFLPHP 323
Oy 330 LDLSQKAFKAIPSTIGVNNHECGFLP-----MKEAPEILSGSNKSLALHL 376
Db 324 QELWASKDFHFPVPSITGINDBYGMILPTIMDPACKIEITRKILPAVL---KSTALKM 379
Oy 377 IONILHIPPQYIALVANEYFDDKSLTEIRDSLDLGLDVEFVVPAITAYHRDAGAPV 436
Db 380 M-----LPBEGDILMEYMGDTEDPETLQAPFEMKDFMFVLPALQVAHFQR-SHAPV 433
Oy 437 YFEFERHPPQCEBDTKPAFVADHADVRFVFGAFLKGLDIWMEGATEEKLSSRQMK 496
Db 434 YFEFERHPPQCEBDTKPAFVADHADVRFVFGAFLKGLDIWMEGATEEKLSSRQMK 489
Oy 497 YMATFARTGNNGNDLSLMPAYNLTEOYLQDLNMSLQORLKEPVDFTSTIP 550
Db 490 YRANFARHGNNSGGLPYWPMDDHDEQLDIDIPSQGRALKARRLOFTWTLIP 543

RESULT 13
Q6PDB7 PRELIMINARY; PRT; 556 AA.
ID Q6PDB7;
AC Q6PDB7;
DT 05-JUL-2004 (TREMblrel. 27, Created)
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)
DE Hypothetical protein BC015286.

```

GN Name=BC015286;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=FVB/N; TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakeley R.W., Young A.C., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [2]
 RC SEQUENCE FROM N.A.
 RP STRAIN=FVB/N; TISSUE=Colon;
 RC Strausberg R.L.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL; BC058815; AAHS815.1; -.
 DR HSSP; P12377; 1K4Y. -.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR InterPro; IPR002018; Carboxylesterase.
 DR InterPro; IPR00379; Ser_esterase.
 DR Pfam; PF00135; Coesterase; 1.
 DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
 DR Hydrolyase: Hypothetical protein.
 SK SEQUENCE 556 AA, 61927 MW, 58950A33C73BECD CRC64;
 Query Match 41.1%; Score 1266.5; DB 2; Length 556;
 Best Local Similarity 47.1%; Pred. No. 5.3e-88;
 Matches 273; Conservative 76; Mismatches 190; Indels 41; Gaps 12;
 QY 1 MPOGLTSSASQWCFPLI--LQPLIGHQMGKGTGSABEPQNTYKLGMIQKQVTVLGSV 58
 DB 1 MPR--SOMNMMDVLLFGLLLGLGHVO-GQDSPEA-SPIRNTHTGGVRSGLVHVXKDTKA 55
 QY 59 PNVNVLGVPAAPPLGSLRFTNPQASPMWDLREATSVYVNLCLNSEMLLLDQMLKVHY 118
 DB 56 GVHFFLGLIPFAKPPVGPLRFAPEAPBPMWGSVGRCTAHFAMCLN---LGVKKEITGLK 111
 QY 119 PKRGVSDCLYNTIYAPAHADYTSKLPVLVWFGCAFKTGSASITPDGSAALAYEDVLVV 178
 DB 112 PAVSTSDCLYNTIYTPAHAHESGSLNLPVWVWINGGGLVAGMASYDGLLAIEDLVVV 171
 QY 179 VOYRLGIFGFFTTWDQHPAGNMAFKDOVALSWQKRIEFPFGDPSVITFGSSAGALISV 238
 DB 172 IQYRLGVLFSPSTODQDARAGNMGFLDQVALKRMVQGNIAHFHGGPPDVITFGSSAGTISV 231
 QY 239 SSLLISPMAGLPHKAIMESGVALIPYLEADHYEKSEDLQVNAFPGNNASDEALLRCL 298
 DB 232 SSHVSVSMXKGLFHGAIMESGVALLPYLITDTSBMS--TTVAKLGGCEAMDEALVRCL 289
 QY 299 RTSPKXELILTSQKTSFTYVNDGAPFPNRPDLISOKAKFAPISITGVNNEHCGFLPM 358
 DB 290 RGSSEBILAIINKLQVMTIPAVNDGEFFPRHPKELSLASEDFHVPVSIIGVNDDEFGWTIP- 348

QY 359 KEAPILSGNSKSLALHIONI-----LHIPPOLYLVANEXFPHDKSLTEIRDS 408
 DB 349 ----VWGSAGQIKETIRENLQAVLNKNTQNLWLPFGCSLLMEIRMGDEDTQIOIQ 403
 QY 409 LLDLLGDVFPVVPALITRYHRDAGAPYFEFPHRPOCEDTYPAFVKADEAVFVF 468
 DB 404 FTEWEDFWFVIALQVAVYFQR-SHASYFEFPHOIASLKDVAFTVKAADHADEIPFVF 462
 QY 469 GGAF--LKGDIWMEGATEEBEKLSSRKMMKYMAFARTGNPNNGMDLSLWPAVYNTEOYLQ 526
 DB 463 GYFFMDKMLD-----TEGKILSRMMKMYAFHAGNPNNSGGLPYFWMDHDEOYLQ 516
 QY 527 LDLMNSIGQRILKEPRVDFWTSTIP--LILASDMLHSPL 563
 DB 517 LDTPAVGRALKSRRLQFWFTLTSQKIQEIRASQDKITEL 556
 RESULT 14
 OCGM54 PRELIMINARY; PRT; 568 AA.
 AC OCGM54;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE LOC443703 protein (Fragment).
 GN Name=LOC443703; (African clawed frog).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [2]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Kidney;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 Dev. Dym. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Klein S., Gernard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
 DR EMBL; BC074230; AAH74230.1; -.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR InterPro: IPR002018; CarboxylaseB.
 DR InterPro: IPR000379; Ser estase.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
 DR Hydrolyase.
 KW HYDROLASE.
 FT NON TER
 SQ SEQUENCE 568 AA; 63032 MW; 9336D08B53931A1 CRC64;

Query Match 41.0%; Score 1263.5; DB 2; Length 568;
 Best Local Similarity 48.2%; Pred. No. 9.2e-88;
 Matches 258; Conservative 87; Mismatches 169; Indels 21; Gaps 8;

QY 28 GKTPSABGPORNRIGWIOGKQVTVGSVPVNVVFLGVPPAPPLGSLRTNPOPAFPM 87
 DB 28 GIYTEDAPRLITTYNGQLGKTVGAKETDRLIHVFKVPAPKAPPIGLKREDFQPEP 87
 QY 88 DNLREATSPYMLCLQNSEML--LLDQHLKVHYPKFVSEDCLYLNTIYAPAHADTGSKL 145
 DB 88 SSIREATENPMWGLQDKKMGQLADFFAKDFPP--VSEDCLYLNTVPTPADRGNEPLP 145
 QY 146 VLWVPGGAFRTGSASIFDGSALAYEDVLYVVVQYRIGIFGFTTWDQHPGKMAFKDQ 205
 DB 146 VVWVFIHGGLTMGAGMGFEGSALAYENVVVSIOYRLGIMGFFSTGDKKARGNYGFLDQ 205
 QY 206 VVALSWYQKNIPEFGDPSVTIFGESAGATSVSSLLSPMAKGLFKHAIMESGVALIPY 265
 DB 206 VVALRWVRDNIKDGAPQSVITIFGESAGLSVSAQVISPUSKGFPHAISSEGVALLPG 265
 QY 266 LEADHYKSEDLQVAVAFCGNNAADS--BALRLCRTKPSKELITLSQTK--SEFTRVVDG 322
 DB 266 LMA---SKTEKILILHVVNAISSCSVSLADCLKKTEDEIIVAIISAMKVAIPAVIDG 322
 QY 323 AFPPNEPLDLISQAFKAIPSIIGVNNHCEGFLLPKKAEPILSGNSKSLALHLIQNLH 382
 DB 323 VFLPKPAEELIASKSNPVPFLIGVNNHFGMILPLALN---ISGYREGMEKDIQSTILV 379
 QY 383 IIPQYLA-----LVANVFPHDKSLTEIRDSLLDLDGVFPVVPALITARYRAGAP 435
 DB 380 ALP-FVHSFTSVVPFIMEEYFGDNDPKELNNPLDVGDIIFVYIPALRTKHYRDSHP 438
 QY 436 VYFEPFHRPQCFEDTPAFVKAHADAEVRVFGGAFPLKGDIVMFEGETEESKLSRKM 495
 DB 439 VYFEPFHRPMSYDSKDDPKADHGDDELIVVGGPFLKSGILFRKSNTEESKLSKIM 498
 QY 496 KYWATFARTGNPNNDLSLMPAYNLTEQYQLDINMSIQRLKEPRVDFTSTIP 550
 DB 499 KYWAFARNGDPNGIGLAEWPKYDEDEYLEIKLQESSQRLKGGRFKFWTVTLIP 553

RESULT 15
 070177 PRELIMINARY; PRT; 561 AA.

AC 070177;
 DT 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Carboxylesterase precursor (EC 3.1.1.1).
 GN Name=carboxylesterase;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wiistar; TISSUE=small intestine;
 RA Score T., Kunitomo T., Isobe M.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 DR EMBL; AB010632; BAA55691.1; -.
 DR HSSP; P37967; IOE3.
 DR GO; GO:0004091; F:carboxylesterase activity; IEA.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.

DR InterPro: IPR002018; CarboxylaseB.
 DR InterPro: IPR000379; Ser estase.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
 DR Hydrolyase; Signal.
 FT SIGNAL
 FT CHAIN
 SQ SEQUENCE 561 AA; 62239 MW; 73A468C33F969398 CRC64;

Query Match 41.0%; Score 1262.5; DB 2; Length 561;
 Best Local Similarity 48.8%; Pred. No. 1.1e-87;
 Matches 275; Conservative 77; Mismatches 179; Indels 33; Gaps 13;

QY 1 MPQG-LTSSASQWKEFLIQLPGLHRQWKGKTPSABEGQRTNRIQKQVTVLGSVVP 59
 DB 1 MPNQLSHMLNVAVLFGLL--LLIHVO-GQDSPSS-SPIRTHTHGQVQKLDHVDYTAG 56
 QY 60 VNVFLGVPPAPPLGSLRTNPOPAFPMNDLREATSPYMLCLQNSEML--LLDQHLK 115
 DB 57 VHTLGIIPFAKPPGPFARFAPPEPPEPMSGVADATSQPMCLQDLIDEGGLDMKMI 115
 QY 116 VHYPKFVSEDCLYLNTIYAPAHADTGSKLPLYVWVPGGAFRTGSASIFDGSALAYEDV 175
 DB 116 --LSSISWSEDCLYLNTIYAPAHAREGSNLPVWVWIHGGLVVGMAWMDGSLITVNEDLV 173
 QY 176 VVVVQYRLGIFGFTTWDQHPGKMAFKDQVVALSWYQKNIPEFGDPSVTIFGESAGA 235
 DB 174 VVTIQYRLGVYGFSTDEHARGMWGYLDQVVALRWQNLIAHFGNPNRYITIFGESAG 233
 QY 236 ISVSSLLSPMAKGLFKHAIMESGVALIPYLEADHYKSEDLQ--VNAHFCGNNAADSSEAL 294
 DB 234 TVSSVHVISPMGQSLFPGAIMESGVALIPDLIS--ETSEVSTTVAKLSCSEAMDSBAL 290
 QY 295 LRCLRTKPSKELITLSQKTSFTRVVDGAPEPNEPLDLSQAKFAIPSIIGVNNHCEGF 354
 DB 291 VRCLRAKSGAELIVNVKFKMIPAVVDGEFLPRPKELIASEDFHPVPSIIIGVNTDEYCC 350
 QY 355 LLPM-----KAPRILSGNSKSLALHLIQNLHLPQYLHVNANVFPHDKSLTEIR 406
 DB 351 TIIPWMTAQIIXE-----LSRENLAVALKDTAAQMLPPECGDILMEBYMNTDSDOTLQ 406
 QY 407 DSLDLGADVFPVVPALITARYHADAGAPVYFESFRRPQCFEDTPAFVKAHADAEVR 466
 DB 407 IQYTEMGDFLFLVLPALQVAHFOR-SHAPVYFYFQIAPSFYKVRPRPHVADADAEV 465
 QY 467 VFGGAFPLKGDIVMEGATEESKLSRKMXYWATFARTGNPNNDLSLMPAYNLTEQY 526
 DB 466 VF-GSFPFGMKLDE--TEERLLSRMKKYNWAFACQGNSEGLPYWPALDHDEGYLQ 521

QY 527 LDINMSIQRLKEPRVDFTSTIP 550
 DB 522 LDTHPAVDRAKARLQFWTKTLP 545

Search completed: June 15, 2005, 09:39:21
 Job time : 95 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 09:25:21 ; Search time 25 Seconds
(without alignments)
2236.078 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079

Sequence: 1 MPGLTSSASQWCFLLQF.....PLSSLTPLSLQPPFFCAP 581

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1298.5	42.2	557	2	A47162 cholinesterase B (E
2	1231.5	40.0	532	2	A43329 60K esterase (E 3
3	1230	39.9	559	1	JC5408 carboxylesterase (
4	1199	38.9	561	2	S47655 carboxylesterase (
5	1186.5	38.5	554	1	S34607 carboxylesterase (
6	1163	37.8	554	2	A39060 carboxylesterase (
7	1162.5	37.8	566	2	S19307 carboxylesterase (
8	1144	37.2	549	2	JX0054 carboxylesterase (
9	1140	37.0	567	1	A41010 carboxylesterase (
10	1139	37.0	561	2	S62788 carboxylesterase (
11	1138.5	37.0	562	2	A55281 carboxylesterase (
12	1136	36.9	540	2	A31584 carboxylesterase (
13	1129.5	36.7	561	2	S71597 carboxylesterase (
14	1123	36.5	561	2	JC2447 carboxylesterase (
15	1121.5	36.4	565	2	S10367 carboxylesterase (
16	1034.5	33.6	539	2	A29923 carboxylesterase (
17	777	25.2	596	1	ACRYE acetylcholinestera
18	764.5	24.8	614	2	A39256 acetylcholinestera
19	763	24.8	614	2	JH0314 acetylcholinestera
20	756	24.6	599	1	A38868 acetylcholinestera
21	754	24.5	614	2	JH0811 acetylcholinestera
22	752	24.4	602	1	ACHU cholinesterase (E 3
23	740	24.0	603	2	S70849 cholinesterase (E 3
24	732	23.8	581	2	C39768 cholinesterase (E 3
25	726.5	23.6	584	2	S46724 acetylcholinestera
26	710.5	23.1	583	2	S10712 acetylcholinestera
27	692.5	22.5	620	2	A54413 triacylglycerol 11
28	660	21.4	745	2	S13586 para-nitrobenzyl e
29	658.5	21.4	489	2	B69680

30	643.5	20.9	597	2	A33668 sterol esterase (E
31	633.5	20.6	599	2	A57701 sterol esterase (E
32	627.5	20.4	664	2	JC7990 acetylcholinestera
33	618	20.1	612	2	A34967 sterol esterase (E
34	594.5	19.3	691	2	JB0150 acetylcholinestera
35	593	19.3	550	1	A34576 crystal protein pr
36	586	19.0	767	2	S47639 acetylcholinestera
37	581.5	18.9	746	2	A25363 acetylcholinestera
38	575	18.6	540	2	S51043 carboxylesterase (
39	572.5	18.6	602	2	T37254 acetylcholinestera
40	570.5	18.5	637	2	S66236 acetylcholinestera
41	569.5	18.5	554	2	T31783 hypochelical prote
42	569.5	18.5	557	2	A56690 esterase - Caenorh
43	562.5	18.3	545	2	A89046 protein B0238.1 [1
44	557.5	18.1	593	1	S25062 triacylglycerol 11
45	551	17.9	562	2	S27782 esterase precursor

ALIGNMENTS

RESULT 1

A47162 cholinesterase B (E 3, - - - -) precursor - mallard

C.Species: Anas platyrhynchos (mallard)

C.Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004

C.Accession: A47162

R.Hwang, C.S.; Kolatukudy, P.E.

J. Biol. Chem. 268, 14278-14284, 1993

A.Title: Molecular cloning and sequencing of cholinesterase B cDNA and stimulation of expe

A.Reference number: A47162; MUID:93300823; PMID:8314791

A.Accession: A47162

A>Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-557 <HMA>

A.Cross-references: UNIPROT:004791; GB:J05493; NID:9213100; PIDN:AAA49223.1; PID:9213101

C.Superfamily: cholinesterase; cholinesterase homology

C.Keywords: hydrolase

F:56-545/Domain: cholinesterase homology <CHE>

Query Match	42.2%	Score 1298.5;	DB 2;	Length 557;
Best Local Similarity	48.4%	Pred. No. 4.1e-94;		
Matches 265;	Conservative 79;	Mismatches 177;	Indels 27;	Gaps 8;
QY	30	TGSAEGQBNTRGCMIGKQVTVGSPVNVNFGVPAAPGLSLEFTNPQASPMDN 69		
DB	23	IGQKAEQBNVTNYSVGRYQVKNAAERSVNVFGLFPAPVGPPLPSEQPEPMKG 82		
QY	90	LRATSYPNLCLQNSEWMLDQHML-----KVHYKPFVSEDCLYLNTYAPAHADTGSKL 144		
DB	83	VRDASVPPKCLQDK---VLQGLSDALITNREKRRLOISBDCLYLNTYTPVSTEQSKL 139		
QY	145	PVLVWFGAPKRTGSASIFDGSALAAVEDVLVVVQYRLGIFGFTTWDQAHGNAFKD 204		
DB	140	PLFVWTHGGGLVSGAASDYGSALAAPNVVVVTVIQLYGLAGYSTGDKARGMWGYLD 199		
QY	205	QVALSWQKNIHFPGSPSSYITFGESAGALSVSGLSLSPAKLFLFKALMESVAL-I 263		
DB	200	QVALQWIOENIIFRRGPGSVITIFGESAGGSVALVLSPLAKLFLFKALSESSTAVRI 259		
QY	264	PYLEAHVYKESBDLQVVAHFGGNNADEALLRCRTKPSKEL--LTLSQKTSFTRVVD 321		
DB	260	LFTE---QPEEQARILAAAGCESSAALVECLREKTBAEMBITLKMPPMFLSASLD 315		
QY	332	GAFPENRDLIDQKAFKAPISIIIGVNNHBCGFLAP-MKEAPELISGNSKSLATLQNI 380		
DB	316	GVFPFKSPRLISSEKVINAVPYIIGVNNCEFGMILPRMKKPEFEGLEKQVAAQVQLST 375		
QY	381	LHI-----PQYLIHVNANRYHDKSLTEIRDSLIDLDGVPFVVVPAITARYHDAQAP 435		
DB	376	LALSPKAPSDIIVDLVYNEYIGVANNRAQVBDGLDLSIADPLFVFSAAVEVAHHHDAQNP 435		
QY	436	VYFVFRFRRPQCFEDTKPAFVKADHAEVRFVFGAFPKGDIVMPEGATEEBKLSRKQM 495		

```

RESULT 3
JC5408
carboxylesterase (EC 3.1.1.1) - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 09-Jul-2004
C:Accession: JC5408
R:Schwartz, H.; Langman, T.; Daig, R.; Becker, A.; Aslanidis, C.; Schmitz, G.
Biochem. Biophys. Res. Commun. 233, 117-120, 1997
A:Title: Molecular cloning and characterization of a novel putative carboxylesterase, pr
A:Reference number: JC5408; MUID:97289502; PMID:9144407
A:Accession: JC5408
A:Molecule type: mRNA
A:Residues: 1-559 <SCH>
A:Cross-references: UNIPROT:O00748; GB:Y09616; NID:G2058317; PIDW:CAA70831.1; PID:G205831
A:Experimental source: intestine
C:Comment: This enzyme hydrolyzes many xenobiotics, such as carboxyl esters, thioesters
C:Gene(s):
A:Gene: GDB:CES2; ICE; CE2
A:Cross-references: GDB:9959011
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein
F:58-544/Domain: cholinesterase homology <CHE>
F:15-95,123-280,291-428/Disulfide bonds: #status predicted
F:111,276/Binding site: carbonylate (Aam) (covalent) #status predicted
F:128,457/Active site: Ser, His #status predicted

Query Match      39.9%; Score 1230; DB 1; Length 559;
Best Local Similarity 46.7%; Pred. No. 1e-88;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

```

Db GDSDSAPRRTHTGTGVLGSLVHVKGANAGVQTEFLGIPFAKPLGLPRFAPEEPESMSGV 95

Qy REATSYPNLCTON-----SEMLLDQNMKMHYPRKGVSEDDCLYNIITYAPAHADTSKLP 145

Db RDGTHHPAKCLDDLAIVESSEFLV---SQPNMTFPPSDMSBEDCLYLSITYPAHSHESNLP 141

Qy VLWMPGGAFKTSASIFDGSALAAVEDLVVVVQVRLGTFGFFTTWDDHAPGNMAFKQ 205

Db VVVVTHGGALVPGMASLYDGSMTLALLENVVYIIQYRLGVLGFFSTGDKHATGNMGYLDQ 201

Qy VVALSWVQKITEFFGGDDSSVTTFEESAGASIVSSLIISPMKGLFHKAIMESGVAIIPY 265

Db VAAERWVQONIAHFEGGNPDRAVTIFGESAGTSVSSIVSGFLFHGAIMESGVALLP 261

Qy LEAHDYKESDL--QVVAHFGCGNMSDSEALLRCITRPSKGLTLISQTKKFTTRVYDGA 323

Db LIA-----SSAVISTVVAHLSACDQVDSBALVGCIRGKSKKEILAIINKPKNIPGVVADGV 317

Qy FEPNEPLDLISOKAFKAIPISIIIVANNHCEGFLP---MKERPEIISGSNKSIALHLIGN 379

Db FLPHRPGQLASADPQRPVPSIVGVANNFEMWIPKMMIYDQKEMDRASQALQAKMLT 377

Qy ILIHPRQYLIHVVANEYFHDKHSLEIRDSLIDLGLGVEFVVVPALITARYHRDAGAPVY 439

Db LLMPEPTFGDLREBYIIGDNGDPQTLQAQOFQEMADSMFVILPALGVAFH-QCSRAIPVY 436

Qy EFRHRPQCFEDTKAPVYADHADAEVRPVFGGAFLLKGLDVMBEGATREBEKLSRKNMKTYA 499

Db EFGQAPSMLKIRPFAKADHGDEL.PVVF-RSFGGANYIKF---TEEBEOLSRKNMKTYA 492

Qy TFAATGPNNGDLSLMPAYNLTEQYLQDLNMSIGORLKEPVRVDPMWTSTIP 550

Db NFANNGPNNGGLPFWPLFDQEBQYLQNLQPAVGRALKAKRLQPMKALP 543

RESULT 4

547655

cardoxy]esterase (EC 3.1.1.1) precursor - golden hamster

Species: Mesocricetus auratus (golden hamster)

R: Stone, T.; Isobe, M.; Takabatake, E.; Wang, C.Y.
Biochim. Biophys. Acta 1207, 138-142, 1994
A: Title: Cloning and sequence analysis of a hamster liver cDNA encoding a novel putative
A: Reference number: S47655; MUID:94318665; PMID:8043605
A: Accession: S47655
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-561 <SON>
C: Cross-references: UNIPROT:O64419; EMBL:D28566; NID:G531238; PIDN:BA05913.1; PID:G531212
C: Superfamily: cholinesterase; cholinesterase homology
C: Keywords: carboxylic ester hydrolase
F: 528-546/Domain: cholinesterase homology <CH>
F: 227,459/Active site: Ser, His #status predicted

Query Match 38.9%; Score 1199; DB 2; Length 561;
Best Local Similarity 47.3%; Pred. No. 2.8e-86;
Matches 252; Conservative 74; Mismatches 181; Indels 26; Gaps 7;

QY 31 GPSAEGPQNRTRLGWIGQKQVTVLGSVPVAVFLGVPAAPPLGSLAFTNPQAPSPNDL 90
DB 26 GQDVSPIRNRHTGQVGRKLVYVKGTVGVAFLGIPAKPVPVGLFAPPEPEPESGV 85
QY 91 REARSYPLCLQNSEBMLLDQHLKMHKPKGVSEDCLYNITYPAHADGSKLPVLYWF 150
DB 86 KDTGISEPMLCQTDPMPKPOISKERKILPTISMSBDCLYNITYPAHMHSSNLPVWMI 145
QY 151 PGCAFPTGSASIPFGSALAAVEDLVVVVQYRLGIFGFFTTMDQADGNAFQKQVALS 210
DB 146 HGGALWGMAMNDGSLAATEDIVISIQYRLGILGFSGDHAGNMGYLDQVALH 205
QY 211 WVKQNIIEFGSDPSVTIIEGSAGAIYSVSLIISPMKGLFHKAIMESGVAIIPYLEAH 270
DB 206 WVOQNINISFGNPGQVTFIIGVSAGTSVSLVSPMSKGLFHGAIMSGVALLPDL-1SD 264
QY 271 YKSEDDQVVAHFGGNASDEALLRCIRTPSKSELLTSGKTSFTRVYDGAFFPPEPL 330
DB 265 TPEAVTPPVANNGSGCEKQSEALVHCLREKTEARILLAINQVFTMPDGVDTLPFRHQ 324
QY 331 DLSQAKFAKPAIPSIYGVNNEHCGFLPM-----KEAPEILSGNSKSLALHIONIL 381
DB 325 ELIASYDFHPVPSIIGVSDSCGNGVPLFMGLDVIKINIRETLPALPKSAEHMM---- 380
QY 382 HIPQYIHLVANEYFHDKSLTEIRDSILLDLGVPFVVPALITARYHRDAGAVYEEF 441
DB 381 -LPRECSDLMLQGYMGDVEDPQTLQAGREBLMKQFMFVIALKVAAYFOR-SHAPVYEEF 438
QY 442 RHRQCF--EETKAPAKADADAEVRFVFGGAF--LKQDIYMGATBEELSLRKMKY 497
DB 439 QHQSFFIKKQKARPSHRADHDGVAFAVFGSDFMGLKIDL-----TEEBQLMKRMKY 492
QY 498 WATPRTGPNPGNDLSLWPAVNLTEQVYLODLNNSIGORLKEPVPDFTSTIP 550
DB 493 WANPRKGNENPSEGFLPYWPELVHDDQYLDQIOPAVGALKSRKLDHFTKILP 545

RESULT 5
S34607
C: carboxylesterase (EC 3.1.1.1) - mouse
C: Species: Mus musculus (house mouse)
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C: Accession: S34607
R: Alda, K.; Moore, R.; Negishi, M.
Biochim. Biophys. Acta 1174, 72-74, 1993
A: Title: Cloning and nucleotide sequence of a novel, male-predominant carboxylesterase
A: Reference number: S34607; MUID:9332638; PMID:7916639
A: Accession: S34607
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-554 <AID>
C: Cross-references: UNIPROT:O63880; GB:S64130; NID:g404388; PIDN:ABD27606.1; PID:g404388
C: Superfamily: cholinesterase; cholinesterase homology
C: Keywords: carboxylic ester hydrolase
F: 46-536/Domain: cholinesterase homology <CH>

F:215_443/Active site: Ser, His #status predicted

Query Match 38.5%; Score 1166.5; DB 1; Length 554;
Best Local Similarity 46.5%; Pred. No. 2,7e-85;
Matches 256; Conservative 81; Mismatches 183; Indels 31; Gaps 10;

DY 13 CFFLLQLPLGLGRQMGKTGPSAEGPQRNRLGIIQGKQVTVLGSPPVVNNFLGIIPFAAP 72
|||
:
DB 3 CLLLPPTTV-----IGPKVTQPEVDTPDLGVRRQGVCKOTDRMWNVFLGIIPFAOP 55
|||
:
DY 73 LGSLEFTNPQPASPMNDLBEATSYPNLCIONSEWMLLDQRLKVHYRKPFGVSDEDCVLNI 132
|||
:
DB 56 LGPLPLFSADLPQQPEGVSDASINPMCLODVERMSNRRTTLBKAKIIFPISECTLTNI 115
|||
:
DY 133 YAPAHADTGSKLPLVWVPFPGAFTKGSASI PDGSALAAVEDVLVWVQYRLGIRGFPTTW 192
|||
:
DB 116 YSPETITGDKKRPVWWIHGGSLRVGSSSTSHDSALAAVGDDVVVTVQYRLGIRGISTG 175
|||
:
DY 193 DGHAFGNMAFKQOVAALSVQKNIEFFGDPSSVTTFGEBSAGAI SVSSLISPAKGLFH 252
|||
:
DB 176 DKHMGNGSGFLDVVAALRWVQGNIAFPFGDPMCVTTFGNSAGGIIVSSLLSPMSAGLFH 235
|||
:
DY 253 KAIMGSAVI IYLEAHDIKESDELQVNAH--FCGNNASPSALLRCTARTKPSKELTLS 310
|||
:
DB 236 RAISOGVAISKILE--DINANSEAQNFNANSVACG--SASPAB--LVQCCLQKGKOLTIKK 291
|||
:
DY 311 QKTSGFTRVVDAFPENEDBLDLSOKAFKAIPSIIIGNNHCEGFLT-----PMKE-A 361
|||
:
DB 292 NVNISYT--VNDSFPQRPOKILLANKQFPVPLFLDTVTHNEFGMLLKPFNNIIDMKHLS 349
|||
:
DY 362 PEILSGSKSLALHLIQNLITHIPPQYLHVANEYPHDKHSLTEIRDSLDDLGLGVFYVP 421
|||
:
DB 350 QEDLLSNSRPLLAH-----MOLPRBIMPVVIDEYLNGSDESAATRYVALQELIGDITVIP 404
|||
:
DY 422 ALITARYHRDACAIFYEFERHRPOCFEDTKPAFYKADHADDEVRFVVGSAFL--KGDIVM 479
|||
:
DB 405 TLIFSKYLODACCPVLFYEHQHTPSSFAPKFPMPVADHSSENAFVGGPFLTDSESLHA 464
|||
:
DY 480 FEGATREEEKLLSRKMVKYATPARTSNPNGNDLSLPAYNLTEQYQYOLDNMISGORLKE 539
|||
:
DB 465 PFEALREBEKQLSTMTMAQWSQFARICNPNGKGLPPRPQLNQLEYLEIGLEPRGVTLKK 524
|||
:
DY 540 PRVDFWTSTIP 550
|||
:
DB 525 GRLEGFWTEPLP 535
|||
:
RESULT 6
A39060 carboxylesterase (RC 3.1.1.1) precursor - mouse
C.Species: Mus musculus (house mouse)
C.Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 09-Jul-2004
C.Accession: A39060
R.Omics: M.; Tepegman, K.; Medda, S.; Elliott, R.W.; Stephenson, D.A.; Grant, S.G.; Ganon
Genomics 9, 344-354, 1991
A.Title: Characterization of a murine cDNA encoding a member of the carboxylesterase multi
A.Reference number: A39060; WUID:91169540; PMID:1840565
A.Accession: A39060
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-554 <OVN>
A.Cross-references: UNIPROT:P23953; GB:M57960; NID:g192853; PID:AAA63297.1; PID:g192854
C.Superfamily: cholinesterase; cholinesterase homology
C.Keywords: carboxylic ester hydrolase
F:50-540/Domain: cholinesterase homology <CH>
F:221,455/Active site: Ser, His #status predicted

Query Match 37.8%; Score 1163; DB 2; Length 554;
Best Local Similarity 44.3%; Pred. No. 1,9e-83;
Matches 241; Conservative 84; Mismatches 193; Indels 26; Gaps 7;

16 LILQPLGLGRQMGKTGPSAEGPQRNRLGIIQGKQVTVLGSPPVVNNFLGIIPFAAPLG 75
|||
:
:


```

Db      11 LAVCPILGH-----SLLPVVDITGQKVLGKYSLSLEGFQPVAVFLGVPKAPPLGS 62
Qy      76 LRFVTPQAPSPMDNLREATSYPNLCLQNSFW--LLDDHMLKXHYKFGVSEDCLYNTLY 133
Db      63 LRFAPQAPSPMDNLREATSYPNLCLQNSFW--LLDDHMLKXHYKFGVSEDCLYNTLY 122
Qy      134 APAAHDTGSKLPVLVWFPFGAFKTSASIFDGSALAAYEDVLVWVVOYRLGIFGFTTMD 193
Db      123 SPADLTFSQSLPVMVWVHGGGLVIGGRSPYNGALSAHENVVYTYQLHIGLWGLFSTGD 182
Qy      194 QHAPGMAKQDVAALSWQKNIIEFGGDPSSVTTFGESAGAISSVSLILSPMAKGLERK 253
Db      183 EHSPPGMALDQIALRWQVDINAFGQNDSTVTTFGESGGSIVSVLVSPLGKDLFHR 242
Qy      254 AIMSSEVAIIPYLEAHDYKESDILQVNAHFCGNMNSDSEALLRCITRKSKLLTLS--- 310
Db      243 AISEGCVNTVNTVNGKKNIOAVNEIATLSQC--NDTSSAAWOCIRKQTESSELTISGKL 300
Qy      311 -QKTSFTRVVDGAFPPNEPDLISQAKFKAIPSIIGVNNHCGFLPM--KEAPE-IL 365
Db      301 VOYNISLSTMIDGVLPKAPBEILAEKSPNTVPYIVGFVKQGFMTIIPMLQNLPECKM 360
Qy      366 SGNNSLMLHLQNLHTPROYLIVANEYFHDKSLTEIRSLDLDLGDVFFVVALIT 425
Db      361 NERTASLLRRHSELNISESMIPAVTEQYLRGVDDPAKKSLLIDMGEDIFFGIPAVIL 420
Qy      426 ARYHDAGAPVYFEFRRHPOCFEDTKPAFVADHAEVRFVFGGAPLKGDIVMFEGATE 485
Db      421 SRSLADAGVSTWYEFRRPSPVSKRPQTVBGDGHDELFFVFGAPLKK-----EGASE 474
Qy      446 EEKLSRKMKKWTAPARTGNPNNGDLSLWPAVNLTEOYLQDLNMSLGRLKEERYDFW 545
Db      475 EETNLSKVMKMANFAIRNGNPGEGLPHPMEYDQEGYLQIGATTQQAQRKAEVAFW 534
Qy      546 TSTI 549
Db      535 TELI 538

```

RESULT 7

```

S19307
C:Species: Sus scrofa domestica (domestic pig)
C/Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C/Accession: S19307, S23607
R:Matsumura, M.; Inoue, H.; Ichinose, M.; Takeda, S.; Miki, K.; Kurokawa, K.; Takahashi, T.
P:Biochem. 1991, 19, 191
A/Title: The nucleotide and deduced amino acid sequences of porcine liver proline-beta-1
A/Reference number: S19307; MUID:92070571; PMID:1959668
A/Accession: S19307
A/Molecule type: mRNA
A/Residues: 1-566 <MAT1>
A/Cross-references: UNIPROT:Q29550; EMBL:X63323; NID:g1930; PID:CAA44929.1; PID:g1931
A/Accession: S23607
A/Molecule type: protein
A/Residues: 19-40 <MAT2>
A/Note: 28-Lys and 33-Leu were also found
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-566/Product: carboxylesterase #status experimental <MAT>
F/51-552/Domain: cholinesterase homology <CHS>
F/80/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/222,467/Active site: Ser, His #status predicted

```

Query Match 37.8%; Score 1162.5; DB 2; Length 566; Best Local Similarity 44.3%; Pred. No. 2,1e-83; Matches 250; Conservative 71; Mismatches 202; Indels 41; Gaps 8;

Qy 12 WCFLLIPLGHRQWKTGPSAEGPQRTLRMTQKQVTVLGSVPVNVVFLGFPAAP 71
 Db 2 WLLPLVLTLSASATW--AGQPASPPVVDPAQGRVIGKYSLSGLGAPVAVFLGVFPAP 59

```

Qy      72 PLGSLRFTNPQAPSPMDNLREATSYPNLCLQNSFW--LLDDHMLKXHYKFGVSEDCLYNTLY 124
Db      60 PLGSLRFTNPQAPSPMDNLREATSYPNLCLQNSFW--LLDDHMLKXHYKFGVSEDCLYNTLY 114
Qy      125 EDCLYNTLYPAHADTGSKLPVLVWFPFGAFKTSASIFDGSALAAYEDVLVWVVOYRLG 184
Db      115 EDCLYNTLYPAHADTGSKLPVLVWFPFGAFKTSASIFDGSALAAYEDVLVWVVOYRLG 174
Qy      185 IGFFTTMDQAPGNMAFKQVAAALSWQKNIIEFGGDPSSVTTFGESAGAISSVSLILS 244
Db      175 IGFFTTMDQAPGNMAFKQVAAALSWQKNIIEFGGDPSSVTTFGESAGAISSVSLILS 234
Qy      245 PMAKLTHKAIMESGVAIIPYLEAHDYKESDILQVNAHFCGNMNSDSEALLRCITRKPSK 304
Db      235 PLAKLTHKAIMESGVAIIPYLEAHDYKESDILQVNAHFCGNMNSDSEALLRCITRKPSK 292
Qy      305 ELTLTQKTSFT-----RVVDGAFPPNEPDLISQAKFKAIPSIIGVNN 349
Db      293 ELTLTQKTSFT-----RVVDGAFPPNEPDLISQAKFKAIPSIIGVNN 352
Qy      350 HECGFLPMKREAPILSGS-NKSLAHLQNLHTPROYLIVANEYFHDKSLTEIRSLDLDLGDVFFVVALIT 405
Db      353 HECGFLPMKREAPILSGS-NKSLAHLQNLHTPROYLIVANEYFHDKSLTEIRSLDLDLGDVFFVVALIT 412
Qy      406 RDSLLDLDLGDVFFVVALITRAHHDAGAPVYFEFRRHPOCFEDTKPAFVADHAEVRFVFGGAPLKGDIVMFEGATE 465
Db      413 RDSLLDLDLGDVFFVVALITRAHHDAGAPVYFEFRRHPOCFEDTKPAFVADHAEVRFVFGGAPLKGDIVMFEGATE 472
Qy      466 FVFGGAPLKGDIVMFEGATEEELSRKMKKWTAPARTGNPNNGDLSLWPAVNLTEOYLQDLNMSLGRLKEERYDFW 525
Db      473 FVFGGAPLKGDIVMFEGATEEELSRKMKKWTAPARTGNPNNGDLSLWPAVNLTEOYLQDLNMSLGRLKEERYDFW 534
Qy      526 QDLNMSLGRLKEERYDFW 549
Db      527 QDLNMSLGRLKEERYDFW 550

```

RESULT 8

```

JX0054
C:Species: Rattus norvegicus (Norway rat)
C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Aug-1998
C/Accession: JX0054
R:Takahashi, Y.; Morohashi, K.; Kawabata, S.; Go, M.; Omura, T.
J:Biochem. 1988, 104, 801-806
A/Title: Molecular cloning and nucleotide sequence of cDNA of microsomal carboxylesterase
A/Reference number: JX0054; MUID:89174514; PMID:3235453
A/Accession: JX0054
A/Molecule type: mRNA
A/Residues: 1-549 <TRX>
A/Experimental source: liver
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; glycoprotein; microsome
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-549/Product: carboxylesterase E1 #status predicted <MAT>
F/50-538/Domain: cholinesterase homology <CHS>
F/79,274,302,375,476/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/221,453/Active site: Ser, His #status predicted

```

Query Match 37.2%; Score 1144; DB 2; Length 549; Best Local Similarity 43.5%; Pred. No. 5.8e-82; Matches 238; Conservative 92; Mismatches 183; Indels 34; Gaps 9;

Qy 16 LILQPLGHRQWKTGPSAEGPQRTLRMTQKQVTVLGSVPVNVVFLGFPAAPPLGS 75
 Db 11 LAVCPILGH-----FSSP-PVVDITGQKVLGKYSLSLEGFQPVAVFLGVPKAPPLGS 62
 Qy 76 LRFVTPQAPSPMDNLREATSYPNLCLQNSFW--LLDDHMLKXHYKFGVSEDCLYNTLY 128
 Db 63 LRFAPQAPSPMDNLREATSYPNLCLQNSFW--LLDDHMLKXHYKFGVSEDCLYNTLY 117
 Qy 129 YLNTYAPAHADTGSKLPVLVWFPFGAFKTSASIFDGSALAAYEDVLVWVVOYRLGIFGF 188

Db 118 YLNIYSPADLTKNRSLPVMWVHGGGLIGGASPYSGALSAHEVVVVTIQLRIGWGL 177
 Qy 189 FTTMDQAPGAWAKDQVAAALSWQKNIIEFGDPSSVTITGESAIAVSLSIPMAK 248
 Db 178 FSTDEHSRGMWALHDOALARWODNIANFGNDSTITIGESAGGVSAVLVSPILAK 237
 Qy 249 GLPKAIKESGVAIIPYLEAHDYKSEDLQVAVHCGNNASDEALRLCRLTKPSKELLT 308
 Db 238 NLFRALISSESVVLTNLIDKQTOAV--AQMIATLISGNTSSAAMVCLQKTEALLE 295
 Qy 309 LSQK--TKSPTRVVDGAFPEPNEPLDLSQKAFKAIPIIIGVNNHCEGLPMPKCAPEILS 366
 Db 296 LTVKLDNTMSGTVIDGVLPKTPPEILTEKSFNVPIYVGNKQDFGMIITPMGNLILSE 355
 Qy 367 G-SKSLALHLIQLI---LHLPQYLIHVAHNYFHDKSLTEIRDSLIDLDLGDVFPVPA 422
 Db 356 GRNMEKMASSFLKFPSPMLNISSEVTPAIIEKYLKGTDDPAKKKELLIDMSDFFGIPA 415
 Qy 423 LITARYHRDAGAPYFYEFRHRPOCFEDTKPAFYKADHADVRFVFGAFPKGDIIVMEEG 482
 Db 416 VLMSRSLRDACAPTYMTEFQYRPSVSDQRQYQGDHGDITFSVYFGPIFK-----EG 469
 Qy 483 ATBEKLLSRKMKYKWTATPARTGNPNNDLSLMPAYNLTEQYQLDNLMSLGQRLKEPRV 542
 Db 470 ASBETMLSKLVKMFMANFANNGNPNBGLPHWPEYDQKEGYLQIGATTQQAQKLGREV 529
 Qy 543 DFWTSTI 549
 Db 530 AFWTELL 536

RESULT 9

A41010
 carboxylesterase (EC 3.1.1.1) precursor, monocyte/macrophage [validated] - human
 N:Alternate names: carboxylesterase, hepatic, monocyte/macrophage serine esterase
 C:Species: Homo sapiens (man)
 C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
 C:Accession: A41010; JH0327; F47376; A49816; PS0280; I61055; A48809; I57004
 R:Kuniger, J.S.; Shi, G.P.; Marx, E.A.; Chin, D.T.; Gerard, C.; Chapman, H.A.
 J. Biol. Chem. 266, 18833-18838, 1991
 A:Title: A serine esterase released by human alveolar macrophages is closely related to
 A:Reference number: A41010; MUID:92011649; PMID:1918003
 A:Accession: A41010
 A:Molecule type: mRNA
 A:Residues: 1-567 <MON>
 A:Cross-references: UNIPROT:P23141; GB:M73499; NID:G179927; PID:AAA35649.1; PID:G179928
 A:Note: parts of this sequence, including the amino end of the mature protein, were conf
 Life Sci. 48, PL43-PL49, 1991
 A:Title: Cloning and sequencing of a human liver carboxylesterase isoenzyme.
 A:Reference number: JH0327; MUID:91148424; PMID:1997784
 A:Accession: JH0327
 A:Molecule type: mRNA
 A:Residues: 61-567 <ION>
 A:Cross-references: GB:M55509; NID:G179929; PID:AAA35650.1; PID:G179930
 R:Shibata, F.; Takagi, Y.; Kitejima, M.; Kuroda, T.; Omura, T.
 Genomics 17, 76-82, 1993
 A:Title: Molecular cloning and characterization of a human carboxylesterase gene.
 A:Reference number: A47376; MUID:94010913; PMID:8406473
 A:Accession: A47376
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA; DNA
 A:Residues: 1-3, 'PALV', 8-11, 'A', 13-567 <SHI>
 A:Cross-references: GB:D21088; NID:G945476; PID:BA04650.1; PID:G9458470
 A:Note: sequence extracted from NCBI backbone (NCBI:P137630) and corrected to correspond
 R:Zschunke, F.; Salmasi, A.; Kteipe, H.; Buck, F.; Parwaresch, M.R.; Radzun, H.J.
 Blood 78, 506-512, 1991
 A:Title: cDNA cloning and characterization of human monocyte/macrophage serine esterase-
 A:Accession: A49816; MUID:91300111; PMID:2070086
 A:Molecule type: mRNA
 A:Residues: 'G', 65-185, 'G', 187-361, 363-567 <ZSC>

A:Cross-references: GB:X52973; NID:G36421; PID:CAA37147.1; PID:G1335304
 R:Riddles, P.W.; Richard, L.J.; Bowles, M.R.; Pond, S.M.
 Gene 108, 289-292, 1991
 A:Title: Cloning and analysis of a cDNA encoding a human liver carboxylesterase.
 A:Reference number: PS0280; MUID:92084150; PMID:1748313
 A:Accession: PS0280
 A:Molecule type: mRNA
 A:Residues: 114, 'R', 116-280, 'A', 282-300, 'IGNSLWYRETOREST', 318-336, 'R', 338-382, 'GSP', 384
 A:Cross-references: GB:M65261; NID:G187028; PID:AAA83932.1; PID:G187029
 A:Experimental source: liver
 A:Note: differences between this sequence and other reports appear to be due to frameshift
 R:Kroetz, D.L.; McBride, O.W.; Gonzalez, F.J.
 Biochemistry 32, 11606-11617, 1993
 A:Title: Glycosylation-dependent activity of baculovirus-expressed human liver carboxylesterase
 A:Reference number: A48809; MUID:94032283; PMID:8218228
 A:Accession: I61055
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-55, 'G', 57-361, 363-535, 'G', 537-567 <KRO1>
 A:Cross-references: GB:L07765; NID:G180949; PID:AAA5711.1; PID:G180950
 A:Accession: A48809
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-17, 'A', 18-55, 'G', 57-535, 'G', 537-567 <KRO2>
 A:Cross-references: GB:L07764; NID:G180947; PID:AAA16036.1; PID:G180948
 A:Genetics:
 A:Gene: GDB:CEB1; HMSR
 A:Cross-references: GDB:128044; OMIM:114835
 A:Map position: 16q13-16q22.1
 C:Superfamily: cholinesterase; cholinesterase homology
 C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; liver
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-145/Domain: propeptide #status predicted <PRO>
 F:150-553/Domain: cholinesterase homology <CHB>
 F:146-567/Product: carboxylesterase #status experimental <MAT>
 F:156-567/Region: endoplasmic reticulum retention signal #status atypical
 F:221,468/Active site: Ser, His #status predicted

Query Match 37.0% Score 1140; DB 1; Length 567;
 Best Local Similarity 44.3% Pred. No. 1.3e-81;
 Matches 248; Conservative 79; Mismatches 197; Indels 36; Gaps 9;

Db 12 WCFPLIPLQLGHRQMKTPSAEGPQNRFLGWIQKQYVVLGSPVAVNVLGVPAP 71
 Qy 2 WLRAPILATLSASAKWH--PSSP-PVVDYTHGVLKFKFSLBGFAPVAFILIPFAKP 58
 Db 72 PLGSLRFTNPQPAPSPMDNLEATSYPNLCLQNSE--WLLDQHLKVAHYPRGVSEDCLY 129
 Qy 59 PLGFLRFTPPQAPAPMSEFVKVATSYPPMCQDPKAGQLSELFTNRKENIPLKLESDCLY 118
 Db 130 LNTYAPAHADTGSCLPVLWFPFGCAFRTGASITDGSNALAYEVVVVVOYRIGIGFF 189
 Qy 119 LNTYTPADLTCKKNLPMWVWHGGGLWGAASYDGLALAAHEVVVVVTIQLRIGWGF 178
 Db 190 TTMDOHAPGNWAFQDQVAAALSWQKNIIEFGDPSSVTITGESAIAVSLSIPMAK 249
 Qy 179 STGBEHRGMWALHDOVAAALRWODNIANSRGNGSTITIGESAGGVSAVLVSPILAK 238
 Db 250 LFRKAIKESGVAIIPYLEAHDYK--SEDLQVAVHCGNNASDEALRLCRLTKPSKELL 307
 Qy 239 LFRHAIKESGVALLSVLVKGDVFLBQALITA---GCKTTTSAVMVHCLRQTEBELT 295
 Db 308 TLSQRTKSFY-----RYVDGAFPEPNEPLDLSQKAFKAIPIIIGVNNHCE 352
 Qy 296 ETTIKMFTSLDLOGDRESQPLIGTYIDGMLTKTPEELQAEENFTVPMVGINKQEF 355
 Db 353 GLIPKMEKAPILSGS--NKSIALHLIQLI---LHLPQYLIHVAHNYFHDKSLTEIRD 407
 Qy 356 GMLTPQQLMSTYPLSEGLDQKATMSLMKSYPLVCIKELIPATRYKLTGTDITVKKD 415
 Db 408 SLDDLGDVFPVPAITARYHRDAGAPYFYEFRHRPOCFEDTKPAFYKADHADVRFV 467
 Qy 416 LFLDLIDVMEGVSVIVARHRDAGAPTYMTEFQYRPSVSDQRQYQGDHGDITFSV 475

199 LHMWDNIATKPGDGSVTITFGESAGSSVSLVLSPLAKNLFORALISSEGVALLTAGLVK 258
 269 HDYKSEDLQVVAHFQGNASDSEALLRCLRTKPSKELLTTSQKRSF----- 316
 259 KOTRPLAKIAVISGCKN--TTSAMVHCLRKOKTEEBELGTTLKNTFLDLHGSROSH 316
 317 ---TRVDGAFPPNEPLDLSQAKFAIPSIIGVNNHCGFLP--MKEAPEILSGSNKSL 372
 317 PFVETVLVDGVALPKRBEILAEKKNFTVPIYVINKQEGFWLPTMMVNPREDVLDQMT 376
 373 ALHLIQN---LHLPPOYLHLVANEYFHDKSLSTEIRDSLLDLGDVFFVPALITARYH 429
 377 AMSLTKKSPFLNLPEDAIIVAAIEKYLKDYTGKNGKQDLLELDVVGVPSPVIVSGH 436
 430 RDAGAPVYFPEFRRHPOCFEDTKPAFVKADHDEVRFVFGCAFLLKGDVMEFGATEBEKL 489
 437 RDAGAPVYFPEFRRHPOCFEDTKPAFVKADHDEVRFVFGCAFLLKGDVMEFGATEBEKL 490
 490 LSRKMKWTATPARTGNGNDLSLMPVNYLTEOYLQDLNMSLGQRLKEPPVDEWTSTI 549
 491 LSRKMKWTATPARTGNGNDLSLMPVNYLTEOYLQDLNMSLGQRLKEPPVDEWTSTI 550

RESULT 12

A:1584
 A:Accession: A1584
 A:Residues: 1-540 <UNP>
 A:Cross-references: UNIPROT:P10959; GB:M20629; GB:X13587; NID:G203279; PIDN:AAA0871.1;
 C:Keywords: carboxylic ester hydrolase; glycoprotein
 F:1-9/Domain: signal sequence #status predicted <SIG>
 F:10-540/Product: carboxylesterase #status predicted <MAT>
 F:41-529/Domain: carboxylesterase homology <CHS>
 F:70,265,266,293,366,467/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:212,444/Active site: Ser, His #status predicted

Query Match 36.9%; Score 1136; DB 2; Length 540;
 Best Local Similarity 43.3%; Pred. No. 2.4e-81;
 Matches 237; Conservative 91; Mismatches 185; Indels 34; Gaps 9;

16 LILQPLIGHQMGKTPGASBGRQNTRLGWIQKQVTVLGSPPVNVFLGVPAPAPLGS 75
 2 LAVCPIMWH-----PSSP--PVVDITKGVKGLKVSLEGFQVAVFLGVFPAPPLGS 53
 76 LRFNTPOPASPMNDLREATSYPNLQI-----SEWLLDQHLKHYPRFVGSBDCL 128
 54 LRFAPPEAEBSWSPKNTTTPRMCSQGVGKLAIDLSTGKESIPLEF-----SEDC 108
 129 YLNIYAPAHADTGSCLPVLVMPFGAFRTGSASTIDGSALAAVEDVTVVVOYRIGIFG 188
 109 YLNIYSPADLTNKSRLPYMWIIGGGLTIIGASPYSGALSAHEVNVVVTTIQRIGFGL 168
 189 FTTMDQAHGMAFPDQVVAALSMVOKNIEFGGDSSTITFGESAGTISVSLISPMAX 248
 169 FSTDEHSRGMWALHDLALRLWQDNLANFGANDSTITFGESAGTISVSLISPLAK 228
 249 GLFPAKIMESGVALIIPYLEADHYEKSEDLQVVAHFQGNASDSEALLRCLRTKPSKELL 308
 229 NLFRAISESGVVLVTNLDKKNQAV--AQMIATISGKNNTSSAMVOCINQKTEALLE 286
 309 LSQK--TKSFRVVDGAFPPNEPLDLSQAKFAIPSIIGVNNHCGFLP--MKEAPEILSG 366
 287 LTVLDNTSMSTVLDGVVLPKTPBEILTEKSFNTVPIYVINKQEGFWLPTMMFVVPDVK 346

367 G-SNKSIALHLIQN---LHLPPOYLHLVANEYFHDKSLSTEIRDSLLDLGDVFFVPA 422
 347 GRMKNKMSFLKRSRPNLNTISESVIIPALIEKYLNGTDPACKKELLDMFSDVFGLRA 406
 423 LITARYHDGAPVYFPEFRRHPOCFEDTKPAFVKADHDEVRFVFGCAFLLKGDVMEFG 482
 407 VLMSRLSDAGAPVYMEFYQRPSPVSDQRPQTVGSDHDEIFSVFTRPFLK-----EG 460
 483 ATEEBKLSRKMKWTATPARTGNGNDLSLMPVNYLTEOYLQDLNMSLGQRLKEPPV 542
 461 ASEEBKLSRKMKWTATPARTGNGNDLSLMPVNYLTEOYLQDLNMSLGQRLKEPPV 520
 543 DFWTSTI 549
 521 AFWTELL 527

RESULT 13

A:1597
 A:Accession: S71597
 A:Residues: 1-561 <YAN>
 A:Experimental source: liver; endoplasmic reticulum
 A:Function: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
 C:Superfamily: cholinesterase; cholinesterase homology
 C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-561/Product: carboxylesterase #status predicted <MAT>
 F:50-551/Domain: cholinesterase homology <CHS>
 F:558-561/Region: endoplasmic reticulum retention signal
 F:79,301/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:221,466/Active site: Ser, His #status predicted

Query Match 36.7%; Score 1129.5; DB 2; Length 561;
 Best Local Similarity 43.4%; Pred. No. 8.3e-81;
 Matches 236; Conservative 83; Mismatches 188; Indels 37; Gaps 9;

31 GPSABGPQNTRLGWIQKQVTVLGSPPVNVFLGVPAPAPLGSILRFTNPOASPMNDL 90
 18 GNPSPPVVDITKGVKGLKVSLEGFQVAVFLGVFPAPAPLGSILRFPAPAPLGS 77
 91 REATSYPNLQI-----LILDQHLKHYPRFVGSBDCLYLNIIYAPAHADTGSCLP 145
 78 KNTTTPRMCSQGVGKLAIDLSTGKESIPLEF-----SEDC 134
 146 VLVMPFGAFRTGSASTIDGSALAAVEDVTVVVOYRIGIFGFTTMDQAHGMAFPDQ 205
 135 VMWVIIGGGLTIIGASPYSGALSAHEVNVVVTTIQRIGIFGFTTMDQAHGMAFPDQ 194
 206 VVALSMVOKNIEFGGDSSTITFGESAGTISVSLISPMAXGLFPAKIMESGVALIIPY 265
 195 VVALHMDQDNLANFGANDSTITFGESAGTISVSLISPMAXGLFPAKIMESGVALIIPY 254
 266 LEADHYEKSEDLQVVAHFQGNASDSEALLRCLRTKPSKELLTTSQK-----TK 314
 255 LFTKQVRA--AKQIADWAGCKTTTSAIIVHCLRKOKTEEBELTEMEKNLILKLSQORDTK 312
 315 S-----FTRVDGAFPPNEPLDLSQAKFAIPSIIGVNNHCGFLP--MKEAPEILSG 368
 313 ESYHFLSTVIDDVVLPKTPBEILTEKSFNTVPIYVINKQEGFWLPTMMFVVPDVK-L 371
 369 NKSIALHLIQ---LHLPPOYLHLVANEYFHDKSLSTEIRDSLLDLGDVFFVPA 425

```

Db      372 DKQWIMLEKFASTIGIPEDIIIVALEKRRKSDDPKIRKRDILAFIGDVFICISWV 431
Qy      426 ARHRAAGAVFYEEYRRHRPQCEDTKPAFVKADHADVRFVFGAFLKGDIMFEGATE 485
Db      432 SRHRAGAPVITYEYQYPSFSSPQRPKDVGDHADVSVFGADILR-----DGASE 485
Qy      486 EEKTLERKMKKYATFARTGNPNNDLSLMPAYNLTEOYLQDLNMSLQORLKEPRVDW 545
Db      486 EEKTLERKMKKYATFARTGNPNNDLSLMPAYNLTEOYLQDLNMSLQORLKEPRVDW 545
Qy      546 TSTI 549
Db      546 TOLL 549

RESULT 14
JC2447
carboxylesterase (EC 3.1.1.1) ES-3 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-Mar-1995 #sequence _revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: JC2447; S23462
R:Robbi, M.; Beaufay, H.
Biochem. Biophys. Res. Commun. 203, 1404-1411, 1994
A:Title: Cloning and sequencing of rat liver carboxylesterase ES-3 (EGASYN).
A:Reference number: JC2447; MUID:95032008; PMID:7945287
A:Accession: JC2447
A:Molecule type: mRNA
A:Residues: 1-561 <ROB>
A:Cross-references: UNIPROT:063108; GB:X81395; NID:9550146; PIDD:CAA57158.1; PID:9550147
A:Experimental source: liver
R:Medda, S.; Proia, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A:Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
A:Reference number: S23460; MUID:92299008; PMID:1606962
A:Accession: S23462
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 503-554, 'R', 556-561 <MED>
A:Cross-references: EMBL:X65295; NID:957557; PIDD:CAA6390.1; PID:957558
A:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-561/Product: carboxylesterase ES-3 #status predicted <MAT>
F:50-551/Domain: cholinesterase homology <CHES>
F:79,107,489/Binding site: carbonylrate (asn) (covalent) #status predicted
F:221,466/Active site: Ser, His #status predicted

Query Match      36.5%  Score 1123;  DB 2;  Length 561;
Best Local Similarity 42.2%  Pred. No. 2.7e-80;
Matches 245;  Conservative 86;  Mismatches 193;  Indels 56;  Gaps 11;

Qy      13 CFFILLOPLGHRWGKTGSAEGPQRTRLGWIQSKQVTVLGSPPVNVNLGVFPAP 72
Db      2 CLVALILVFLAFTAG--GHPSSLPVVDTLQGVKLGVKYSLEGTQVAVNPLFGSPKRP 59
Qy      73 LGSIRFNPQASWWDILREKTSYFNLCLQ-----SEWLLIDHMLKVHYPKF 121
Db      60 LGSIRFNPQASWWDILREKTSYFNLCLQ-----SEWLLIDHMLKVHYPKF 121
Qy      122 GVSDECCYLYNTPAFAHADTGSKLPLVWFPBGAKRTGSASIFDSALAAEDVLYVVVQY 181
Db      113 --SDDCYLYNITPADTKDRPLRPMVWVHGGGLVLDGASITDGLASTHENVVVVVQY 170
Qy      182 RLGIFFGFTTWDQAPGNMAFKQVAALSWQKNIEFFGDPSSVTIFGSAGASVSSL 241
Db      171 RLGIWGFSTGDEHSRNGMGLDQVALHWWQDINDFGDDPGSVTIFGSAGGESVSL 230
Qy      242 ILSPMAKGLFHKXIMESGVAILPYL-BAHYEKSEDDQVVAHCGNNASSEMLRLCLRT 300
Db      231 VLSPLANLHFKXISGVALTAGLVKKNRPLAEKLAIVS---GCSTYSASMVHCLQ 287
Qy      301 KPSKELLTISQKTSFT-----RVWDGAFPPNBPDLDSQAKAKALPSII 345

```

```

Db      288 KTEBELTFTTKLMLFSLDLHDSROSYPFVPVTLDDGVLPKMEBEILAEKDFVTPYIV 347
Qy      346 GVNNHECGFLP-MKEAPEILSGSKSLAHLION---ILHIPOYLHANEYFHDKS 401
Db      348 GINKEPFWIIPMTANNYPSPMKLDPMATYISLKKSSSLMLPPEALPVAVEKTLRHTDD 407
Qy      402 LTERDLSLDLIGVFFVPPALITARYHRDAGAVFYEEYRRHRPQCEDTKPAFVKADHA 461
Db      408 PDRNKDQLLELIGDVIIEGVPSIVISRGHRDAGARTMYEFQYRPSFSSKKPSTVGDHG 467
Qy      462 DEVRFVFGAFLKGDIMFEGATEEEKLSRKMKKYATFARTGNPNNDLSLMPAYNL 521
Db      468 DEIVSFGAPLTRG-----GTSKEELNLSKMMKFWANFARNGNPNQGLPHMPEYDOK 521
Qy      522 EOYLQDLNMSLQORLKEPRVDFTSTPILISADMLHS 561
Db      522 EGYLIQIGATTQQAOKLKEKXVAFWSE-----LLAKRLHA 556

RESULT 15
S10367
carboxylesterase (EC 3.1.1.1) ES-10 precursor, microsomal - rat
N:Alternate names: hydrolase A
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1991 #sequence _revision 30-Jun-1991 #text_change 09-Jul-2004
C:Accession: S10367; S12468; S1202; S23460; S14361
R:Robbi, M.; Beaufay, H.; Octave, J.N.
Biochem. J. 269, 451-458, 1990
A:Title: Nucleotide sequence of cDNA coding for rat liver pl 6.1 esterase (ES-10), a cart
A:Reference number: S10367; MUID:90351366; PMID:2286485
A:Accession: S10367
A:Molecule type: mRNA
A:Residues: 1-565 <ROB1>
A:Cross-references: UNIPROT:Q9R135; EMBL:X51974
A:Note: 168-Gln, 247-Tyr, 423-Met, and 506-Asn were also found
R:Robbi, M.
submitted to the EMBL Data Library, February 1990
A:Reference number: S12468
A:Accession: S12468
A:Molecule type: protein
A:Residues: 1-264, 'R', 266-565 <ROB2>
A:Cross-references: EMBL:X51974; NID:956898; PIDD:CAA36236.1; PID:956899
R:Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A:Title: Purification and characterization of two rat liver microsomal carboxylesterases
A:Reference number: S51202; MUID:95077430; PMID:7986098
A:Accession: S51202
A:Molecule type: protein
A:Residues: 19-48 <WOR>
R:Medda, S.; Proia, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A:Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting ti
A:Reference number: S23460; MUID:92299008; PMID:1606962
A:Accession: S23460
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-185, 'Q', 187-422, 'W', 424-505, 'N', 507-565 <MED>
A:Cross-references: EMBL:X65296; NID:957553; PIDD:CAA46391.1; PID:957554
R:Gauselad, R.; Sletten, K.; Lovhaug, D.; Fonnum, F.
Biochem. J. 274, 693-697, 1991
A:Title: Purification and characterization of carboxylesterases from rat lung.
A:Reference number: S14361; MUID:91190080; PMID:2012599
A:Accession: S14361
A:Molecule type: protein
A:Residues: 19-26, 'D', 28-37 <GAN>
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-565/Product: carboxylesterase #status predicted <MAT>
F:50-551/Domain: cholinesterase homology <CHES>
F:79,489/Binding site: carbonylrate (asn) (covalent) #status predicted

```

F:221,466/active site: Ser, His #status predicted

Query Match 36.4%; Score 1121.5; DB 2; Length 565;

Best Local Similarity 42.4%; Pred. No. 3.6e-80; Mismatches 245; Conservative 79; Mismatches 199; Indels 55; Gaps 11;

```

QY 12 WCFPLIIOPLIGHQMGKTGPSABGPQNTIRLGMIOGQVTVLGSPPVNVFLGVFPAP 71
DB 8 WLF-----LAACIAGY--PSSP-PVNTYKGVKVLGKYNLEGFAGQVAVFLGIFPAK 58
QY 72 PLGSLRFTNPQASPMWDLREATSYPNLCLQN-----SEWLLIDQMLKHYPKGV 124
DB 59 PLGSLRFPAPQPAEPWNVKNTTSYPMCSODAVGQVLSLFTNRKENIPLQF-----S 113
QY 125 EDCYINIVYARAHADTSGKLPVLVWFPGAFKTSASIFDSALAAVEDLVVVVQYRLG 184
DB 114 EDCYLVNYTPADLTKNRSLPVMWVHGGGLVVGASTYDQVLSAHENVVVTTQYRLG 173
QY 185 IFGFFTWDQAPGNMAFKDQVAALSWYQKNIEFPQDPSVTTIFGSAGAISVSSLIS 244
DB 174 IWGFPSITGDEHSRGMGHLDQVAALHWQDNINANFGNPGSVTTIFGSAGGSVALVLS 233
QY 245 PMAGLFPKAIMESGVALLPYLEAHDYKESBDLOVVAHFCGNNASDSEALLRCLRTKPSK 304
DB 234 PLAKNLFHRAISESGVLTSLALITDSKRIANL--IATLSGCKTTTSAMVYHCLRQKTED 291
QY 305 ELTLISQKTSF-----TRVVDGAFEPNEPLDLSQKAFKAIPIIIGVNN 349
DB 292 ELLETSLKLNLFKLDLGNPKESYPLPTVIDGVVLPKTPBEILAEKSFNTVPYIVGINK 351
QY 350 HECGFLLPMKEAPELLISG-----SNKSLALHLIQNLHIPOYLHLVANEYFHDGHSLTE 404
DB 352 QEFGWIIPTLMGYPLSEKLDQKTAKSLLNKSYPYTLKISEKMIPIVAEKIFGCTDDPAK 410
QY 405 IRDSLILDLGDFVFPVPPALITARYHRDAGAPVYFEPHRRPQCFEDTKPAFYKADHADEV 464
DB 411 RKDLFQDLVAVDVIIGVPSVWVSRSHRDAGAPTFMYEFEPSPFSVAMRPKTVIGDHGDEL 470
QY 465 RFVFGGAPLKGDIVFEGATEEBEKLRSKMKKYATTPARTGNPNGNDSLMPAYNLTEQY 524
DB 471 FSVFSGSPFLK-----DSASEBETNLSKVMKYMANFANRNSPNGGGLPHWPEYDQKEGY 524
QY 525 LQDLNLSIGQRLKEPRVDFTSTIPLIISASDMLHSP 562
DB 525 LKIGASTQAQRLKDKGVAFWSE-----LRKKAABEP 557

```

Search completed: June 15, 2005, 09:39:49
Job time : 27 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 09:39:26 ; Search time 86 Seconds
(without alignments)
2589.736 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079
Sequence: 1 MPOGLTSSASQWCFLLIQP.....PLSSLTFLSLIQPFPPFCAP 581

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum March 0%
Maximum March 10%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10E_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	100.0	581	US-10-023-515-2	Sequence 2, Appl1
2	3079	100.0	581	US-10-674-636-2	Sequence 2, Appl1
3	3079	100.0	581	US-10-757-262-46	Sequence 46, Appl1
4	2915	94.7	575	US-10-451-168-91	Sequence 91, Appl1
5	2900	94.2	642	US-10-433-256-10	Sequence 92, Appl1
6	2607.5	84.7	581	US-10-451-168-92	Sequence 2375, Ap
7	2606	84.6	525	US-10-094-749-2375	Sequence 196, Ap
8	2597.5	84.4	581	US-10-114-270-196	Sequence 2, Appl1
9	2588.5	84.1	618	US-10-381-898-2	Sequence 2219, Ap
10	2454	79.7	469	US-10-104-047-2219	Sequence 2, Appl1
11	1985	64.5	542	US-10-233-933A-2	Sequence 2, Appl1

12	1985	64.5	542	16	US-10-233-933A-4	Sequence 4, Appl1
13	1780	57.8	356	15	US-10-451-168-93	Sequence 99, Appl1
14	1230	39.9	549	9	US-09-925-298-689	Sequence 689, App
15	1230	39.9	549	14	US-10-102-806-689	Sequence 689, App
16	1230	39.9	550	15	US-10-267-756-5	Sequence 5, Appl1
17	1230	39.9	559	16	US-10-858-271-28	Sequence 28, Appl1
18	1186.5	38.5	554	9	US-09-895-860-4	Sequence 4, Appl1
19	1186.5	38.5	554	15	US-10-377-072-4	Sequence 4, Appl1
20	1186.5	38.5	554	16	US-10-377-072-4	Sequence 23, Appl1
21	1179.5	38.3	571	10	US-09-931-836-23	Sequence 1090, Ap
22	1179.5	38.3	571	10	US-09-833-245-1090	Sequence 23, Appl1
23	1179.5	38.3	571	13	US-10-036-342-23	Sequence 23, Appl1
24	1179.5	38.3	571	13	US-10-036-041-23	Sequence 542, App
25	1179.5	38.3	571	14	US-10-028-072-542	Sequence 542, App
26	1179.5	38.3	571	14	US-10-035-855-23	Sequence 23, Appl1
27	1179.5	38.3	571	14	US-10-140-808-542	Sequence 542, App
28	1179.5	38.3	571	14	US-10-121-049-542	Sequence 542, App
29	1179.5	38.3	571	14	US-10-123-904-542	Sequence 542, App
30	1179.5	38.3	571	14	US-10-140-470-542	Sequence 542, App
31	1179.5	38.3	571	14	US-10-175-746-542	Sequence 542, App
32	1179.5	38.3	571	14	US-10-176-918-542	Sequence 542, App
33	1179.5	38.3	571	14	US-10-176-921-542	Sequence 542, App
34	1179.5	38.3	571	14	US-10-227-884-210	Sequence 210, App
35	1179.5	38.3	571	14	US-10-036-214-23	Sequence 23, Appl1
36	1179.5	38.3	571	14	US-10-137-865-542	Sequence 542, App
37	1179.5	38.3	571	14	US-10-140-474-542	Sequence 542, App
38	1179.5	38.3	571	14	US-10-035-719-23	Sequence 23, Appl1
39	1179.5	38.3	571	14	US-10-142-431-542	Sequence 542, App
40	1179.5	38.3	571	14	US-10-143-114-542	Sequence 542, App
41	1179.5	38.3	571	14	US-10-230-163-210	Sequence 210, App
42	1179.5	38.3	571	14	US-10-036-160-23	Sequence 23, Appl1
43	1179.5	38.3	571	14	US-10-230-338-210	Sequence 210, App
44	1179.5	38.3	571	14	US-10-142-419-542	Sequence 542, App
45	1179.5	38.3	571	14	US-10-218-631-210	Sequence 210, App

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Siles-Santiago, Inmaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2

Query Match 100.0%; Score 3079; DB 13; Length 581;
Best Local Similarity 100.0%; Pred. No. 1,1e-277;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPOGLTSSASQWCFLLIQPLIGHROWKGTGSAAGPQNRIRLGIQKQKQVTVLSSPPV 60
Db 1 MPOGLTSSASQWCFLLIQPLIGHROWKGTGSAAGPQNRIRLGIQKQKQVTVLSSPPV 60
QY 61 NVFLGVPAAPPLGSLRFTNPQAPSPMDNLRBATSYPNVLCIONSEMLLDQMLKVHYPK 120

Db 61 NVFLGVPAAPPLGSLRFTNPOPASPDNDLREATSYNCLONSEWILLDQHLKHYRK 120
QY 121 FGVSSECLYLNITYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
Db 121 FGVSSECLYLNITYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
QY 181 YRLGIFGFFTTMDQAPGNMFAFKDQVALSWQKNIIEFGGDPSSVTITFGESAGASVSS 240
Db 181 YRLGIFGFFTTMDQAPGNMFAFKDQVALSWQKNIIEFGGDPSSVTITFGESAGASVSS 240
QY 241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDYKESDLOVVAHFCGNNADESEALLRCURT 300
Db 241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDYKESDLOVVAHFCGNNADESEALLRCURT 300
QY 301 KPSKELLTLOSOKTSFTRVVDGAFPPNEPDLDSQAKFAKIPSIIGVNNHCGFLPMKE 360
Db 301 KPSKELLTLOSOKTSFTRVVDGAFPPNEPDLDSQAKFAKIPSIIGVNNHCGFLPMKE 360
QY 361 APELISGNSKSLAHLIQLNIHLIPQYLHLVANEYFHDGSLTEIRDSLLDLDGDFPVV 420
Db 361 APELISGNSKSLAHLIQLNIHLIPQYLHLVANEYFHDGSLTEIRDSLLDLDGDFPVV 420
QY 421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
Db 421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
QY 481 EGATEEEKLSRRKMKKWTATFARTGNPNNDLSLWPAVNLTEOYLQDLNMSLGQRLEP 540
Db 481 EGATEEEKLSRRKMKKWTATFARTGNPNNDLSLWPAVNLTEOYLQDLNMSLGQRLEP 540
QY 541 RVDFTSTIPLILSASDMLHSPSLSTFLSLQPPFFFCAP 581
Db 541 RVDFTSTIPLILSASDMLHSPSLSTFLSLQPPFFFCAP 581

RESULT 2
US-10-674-636-2
; Sequence 2, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; APPLICANT: Slios-Santiago, Imaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT FILING DATE: US/10/674,636
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-674-636-2

Query Match 100.0%; Score 3079; DB 15; Length 581;
Best Local Similarity 100.0%; Pred. No. 1.le-277;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPOGLTSSASQWCFLLIPLILGHROGKTPSABGQNRRLGWIQKQVTVLGSVPV 60
Db 1 MPOGLTSSASQWCFLLIPLILGHROGKTPSABGQNRRLGWIQKQVTVLGSVPV 60
QY 61 NVFLGVPAAPPLGSLRFTNPOPASPDNDLREATSYNCLONSEWILLDQHLKHYRK 120
Db 61 NVFLGVPAAPPLGSLRFTNPOPASPDNDLREATSYNCLONSEWILLDQHLKHYRK 120

QY 121 FGVSSECLYLNITYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
Db 121 FGVSSECLYLNITYAPAHADTGSCLPVLVWPGGAFKTSASIFDGSALAAVEDVLVVVVQ 180
QY 181 YRLGIFGFFTTMDQAPGNMFAFKDQVALSWQKNIIEFGGDPSSVTITFGESAGASVSS 240
Db 181 YRLGIFGFFTTMDQAPGNMFAFKDQVALSWQKNIIEFGGDPSSVTITFGESAGASVSS 240
QY 241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDYKESDLOVVAHFCGNNADESEALLRCURT 300
Db 241 LILSPMAKGLFHKAIMESGVAIIPYLEAHDYKESDLOVVAHFCGNNADESEALLRCURT 300
QY 301 KPSKELLTLOSOKTSFTRVVDGAFPPNEPDLDSQAKFAKIPSIIGVNNHCGFLPMKE 360
Db 301 KPSKELLTLOSOKTSFTRVVDGAFPPNEPDLDSQAKFAKIPSIIGVNNHCGFLPMKE 360
QY 361 APELISGNSKSLAHLIQLNIHLIPQYLHLVANEYFHDGSLTEIRDSLLDLDGDFPVV 420
Db 361 APELISGNSKSLAHLIQLNIHLIPQYLHLVANEYFHDGSLTEIRDSLLDLDGDFPVV 420
QY 421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
Db 421 PALITARYHRDAGAPVYFYEFRHRPQCFEDTKPAFVADHADEVRFVFGAFLKGDIVMF 480
QY 481 EGATEEEKLSRRKMKKWTATFARTGNPNNDLSLWPAVNLTEOYLQDLNMSLGQRLEP 540
Db 481 EGATEEEKLSRRKMKKWTATFARTGNPNNDLSLWPAVNLTEOYLQDLNMSLGQRLEP 540
QY 541 RVDFTSTIPLILSASDMLHSPSLSTFLSLQPPFFFCAP 581
Db 541 RVDFTSTIPLILSASDMLHSPSLSTFLSLQPPFFFCAP 581

RESULT 3
US-10-757-262-46
; Sequence 46, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karichet, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 5154,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; FILE REFERENCE: MFI03-007PRLNOMNIM
; CURRENT APPLICATION NUMBER: US/10/757,262
; PRIOR FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332

PRIOR FILING DATE: 2003-09-26
 NUMBER OF SEQ ID NOS: 136
 SOFTWARE: FASTSEQ for windows Version 4.0
 SEQ ID NO 46
 LENGTH: 581
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-757-262-46

Query Match 100.0%; Score 3079; DB 16; Length 581;
 Best Local Similarity 100.0%; Pred. No. 1,1e-277;
 Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPOGLTSSASQWCFPLITLQPLIGHKQMGKTPSABGPQNTRLGIQKQVTVLSSPVV 60
 DB 1 MPOGLTSSASQWCFPLITLQPLIGHKQMGKTPSABGPQNTRLGIQKQVTVLSSPVV 60
 QY 61 NVPLGVPFAAPPLGSLRFTNPQASPMNLRATSYPNLCQNSEWLLDDHMLKVHYPK 120
 DB 61 NVPLGVPFAAPPLGSLRFTNPQASPMNLRATSYPNLCQNSEWLLDDHMLKVHYPK 120
 QY 121 FGVSDECLYLYIYAPAHADTGSKLPLYVFPFGAKFTGSASIFDGSALAAVEDVLYVYVQ 180
 DB 121 FGVSDECLYLYIYAPAHADTGSKLPLYVFPFGAKFTGSASIFDGSALAAVEDVLYVYVQ 180
 QY 121 YRLGIFGFTTWDQAPGNMAFKDQVALSWQKNIIEFGDPSVTTFGESAGAISSVS 240
 DB 121 YRLGIFGFTTWDQAPGNMAFKDQVALSWQKNIIEFGDPSVTTFGESAGAISSVS 240
 QY 181 YRLGIFGFTTWDQAPGNMAFKDQVALSWQKNIIEFGDPSVTTFGESAGAISSVS 240
 DB 181 YRLGIFGFTTWDQAPGNMAFKDQVALSWQKNIIEFGDPSVTTFGESAGAISSVS 240
 QY 241 LILSPMAGLPHKAIMESGVALIPLYEADHYKSEBDLQVVAHFCGNNSDSEALRCURT 300
 DB 241 LILSPMAGLPHKAIMESGVALIPLYEADHYKSEBDLQVVAHFCGNNSDSEALRCURT 300
 QY 301 KPSKELLTLOKTKSFTRVVDGAFPPNEPDLISQAKFAIPSIIGVNNHECGFLPMKE 360
 DB 301 KPSKELLTLOKTKSFTRVVDGAFPPNEPDLISQAKFAIPSIIGVNNHECGFLPMKE 360
 QY 361 APELISGNSKSLAHLIQLNIHLIPQYLYLVANNEYFHDGSLTEIRDSLLDLGVPVAV 420
 DB 361 APELISGNSKSLAHLIQLNIHLIPQYLYLVANNEYFHDGSLTEIRDSLLDLGVPVAV 420
 QY 421 PALITARHARDAGAVYFYEFRHRPQCFEDTAPAVKADHAEVRFVFGAFLKGDIVWF 480
 DB 421 PALITARHARDAGAVYFYEFRHRPQCFEDTAPAVKADHAEVRFVFGAFLKGDIVWF 480
 QY 481 EGATEEEKLISRKMKKYATPARTGNPNNDLSLMPAYNLTEQYQLDLNMSLQRLKEP 540
 DB 481 EGATEEEKLISRKMKKYATPARTGNPNNDLSLMPAYNLTEQYQLDLNMSLQRLKEP 540
 QY 541 RVDFTSTITPLILSASDMLHSPSLTFLSLQPPFFFCAP 581
 DB 541 RVDFTSTITPLILSASDMLHSPSLTFLSLQPPFFFCAP 581

RESULT 4
 US-10-451-168-91
 Sequence 91, Application US/10451168
 Publication No. US20040091969A1

GENERAL INFORMATION:
 APPLICANT: SMITHKLINE BEECHAM CORPORATION
 APPLICANT: SMITHKLINE BEECHAM P.L.C.
 APPLICANT: GILKINO GROUP LIMITED
 TITLE OF INVENTION: NOVEL COMPOUNDS
 FILE REFERENCE: GP50039
 CURRENT APPLICATION NUMBER: US/10/451,168
 PRIOR FILING DATE: 2003-11-12
 PRIOR APPLICATION NUMBER: PCT/US01/49232
 PRIOR FILING DATE: 2000-12-17
 PRIOR APPLICATION NUMBER: 60/256,710
 PRIOR FILING DATE: 2000-12-19
 PRIOR APPLICATION NUMBER: 60/257,048
 PRIOR FILING DATE: 2000-12-20
 PRIOR APPLICATION NUMBER: 60/260,482

File 16/90

PRIOR FILING DATE: 2001-01-09
 PRIOR APPLICATION NUMBER: 60/264,922
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: 60/266,797
 PRIOR FILING DATE: 2001-02-06
 PRIOR APPLICATION NUMBER: 60/276,988
 PRIOR FILING DATE: 2001-03-19
 PRIOR APPLICATION NUMBER: 60/281,535
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: 60/289,622
 PRIOR FILING DATE: 2002-06-28
 NUMBER OF SEQ ID NOS: 110
 SOFTWARE: FASTSEQ for windows Version 4.0
 LENGTH: 575
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-451-168-91

Query Match 94.7%; Score 2915; DB 15; Length 575;
 Best Local Similarity 97.0%; Pred. No. 2,1e-262;
 Matches 553; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 12 WCFLLIQLIGHKQMGKTPSABGPQNTRLGIQKQVTVLSSPVVNVFLGVFPAP 71
 DB 14 WAIWLAAP-----TKGPSABGPQNTRLGIQKQVTVLSSPVVNVFLGVFPAP 65
 QY 72 PLGSLRFTNPQASPMNLRATSYPNLCQNSEWLLDDHMLKVHYKFGVSEBCLYLN 131
 DB 66 PLGSLRFTNPQASPMNLRATSYPNLCQNSEWLLDDHMLKVHYKFGVSEBCLYLN 125
 QY 132 IYAPAHADTGSKLPLYVFPFGAFTGSASIFDGSALAAVEDVLYVYVQYRLGIFGFTT 191
 DB 126 IYAPAHADTGSKLPLYVFPFGAFTGSASIFDGSALAAVEDVLYVYVQYRLGIFGFTT 185
 QY 192 WDQAHAGMNAFQDQVALSWQKNIIEFGDPSVTTIFGESAGAISSVSLILSPMAKGLF 251
 DB 186 WDQAHAGMNAFQDQVALSWQKNIIEFGDPSVTTIFGESAGAISSVSLILSPMAKGLF 245
 QY 252 HKAIMESGVALIPLYEADHYKSEBDLQVVAHFCGNNSDSEALRCURTKEKELLTISO 311
 DB 246 HKAIMESGVALIPLYEADHYKSEBDLQVVAHFCGNNSDSEALRCURTKEKELLTISO 305
 QY 312 KTKSFTRVVDGAFPPNEPDLISQAKFAIPSIIGVNNHECGFLPMKEAPEILSGNSKS 371
 DB 306 KTKSFTRVVDGAFPPNEPDLISQAKFAIPSIIGVNNHECGFLPMKEAPEILSGNSKS 365
 QY 372 LALHLIQLNIHLIPQYLYLVANNEYFHDGSLTEIRDSLLDLGDFVFPALITARYARD 431
 DB 366 LALHLIQLNIHLIPQYLYLVANNEYFHDGSLTEIRDSLLDLGDFVFPALITARYARD 425
 QY 432 AGAPVYFYEFRHRPQCFEDTAPAVKADHAEVRFVFGAFLKGDIVMEGATEEEKLS 491
 DB 426 AGAPVYFYEFRHRPQCFEDTAPAVKADHAEVRFVFGAFLKGDIVMEGATEEEKLS 485
 QY 492 RRMKMYATPARTGNPNNDLSLMPAYNLTEQYQLDLNMSLQRLKEPVDFTSTITPL 551
 DB 486 RRMKMYATPARTGNPNNDLSLMPAYNLTEQYQLDLNMSLQRLKEPVDFTSTITPL 545
 QY 552 ILSASDMLHSPSLTFLSLQPPFFFCAP 581
 DB 546 ILSASDMLHSPSLTFLSLQPPFFFCAP 575

X

RESULT 5
 US-10-433-256-10
 Sequence 10, Application US/10433256
 Publication No. US20040081980A1
 GENERAL INFORMATION:
 APPLICANT: SANJUNWALA, Madhusudan M.; YAO, Monique G.
 APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
 APPLICANT: ARVIZU, Chandra S.; RING, Huijun Z.
 APPLICANT: LEE, Ernestine A.; DING, Li

```

/ APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
/ APPLICANT: YUE, Henry; TRIBOULEY, Catherine M.
/ APPLICANT: LU, Dying Aina M.; LAU, Preeti G.
/ APPLICANT: WARREN, Bridget A.; YANG, Junning
/ APPLICANT: CHAMLA, Narinder K.; NGUYEN, Daniel B.
/ APPLICANT: GANDHI, Ameena R.; LU, Yan
/ APPLICANT: ISON, Craig H.
/ TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
/ FILE REFERENCE: PI-0113 USN
/ CURRENT APPLICATION NUMBER: US/10/433,256
/ PRIOR FILING DATE: 2003-05-30
/ PRIOR APPLICATION NUMBER: PCT/US01/47429
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: US 60/254,308
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: US 60/256,189
/ PRIOR FILING DATE: 2000-12-15
/ PRIOR APPLICATION NUMBER: US 60/257,713
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: US 60/262,706
/ PRIOR FILING DATE: 2001-01-19
/ PRIOR APPLICATION NUMBER: US 60/266,020
/ PRIOR FILING DATE: 2001-02-02
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PERL Program
/ SEQ ID NO: 10
/ LENGTH: 642
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No: 6538080CD1
US-10-433-256-10

```

211/29/00

```

Query Match      94.2%; Score 2900; DB 15; Length 642;
Best Local Similarity 99.5%; Pred. No. 6,3e-261;
Matches 548; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 31 GPSAGPQPNTRLGWTIGKQVTVLGSPPVNVFLGPPAPPLGSLRFTNPQAPSPMDNL
DB 92 GPSAGPQPNTRLGWTIGKQVTVLGSPPVNVFLGPPAPPLGSLRFTNPQAPSPMDNL 151
QY 91 REATSYPNLCLONSEWMLLDQMLKVHYPKFGVSEDCLYNTIYAPAHADTGSKLPVYWF 150
DB 152 REATSYPNLCLONSEWMLLDQMLKVHYPKFGVSEDCLYNTIYAPAHADTGSKLPVYWF 211
QY 151 PGGAFTKGSASIFDGSALAAYEDVLVVVQYRLGIFGFTTMDQHPANMAFKQVVALS 210
DB 212 PGGAFTKGSASIFDGSALAAYEDVLVVVQYRLGIFGFTTMDQHPANMAFKQVVALS 271
QY 211 WVOQNIIEFFGDDPSVYTFGFSAGAISSVSLILSPMAKGLFHKAIMESGVALIPLYBAHD 270
DB 272 WVOQNIIEFFGDDPSVYTFGFSAGAISSVSLILSPMAKGLFHKAIMESGVALIPLYBAHD 331
QY 271 YEKSEDIQVVAHFPGNNASDSEALLRCLRTKPSKELTLISOKTSFTFRVVDGAFPPNEPL 330
DB 332 YEKSEDIQVVAHFPGNNASDSEALLRCLRTKPSKELTLISOKTSFTFRVVDGAFPPNEPL 391
QY 331 DLSQKAFKAIPIISIIIGVNNHCEGFLIPMKCAPETILSSNKSALAHLLIONTIHLIPQYTLH 390
DB 392 DLSQKAFKAIPIISIIIGVNNHCEGFLIPMKCAPETILSSNKSALAHLLIONTIHLIPQYTLH 451
QY 391 VANEYFHDKSLTEIRDSLDDLDGDFVFPVVPALITAYHRDAGAPVYFFRRRPOCFED 450
DB 452 VANEYFHDKSLTEIRDSLDDLDGDFVFPVVPALITAYHRDAGAPVYFFRRRPOCFED 511
QY 451 TKFAFYVADHAEVRFVFGAFLKGDIVMEGATEEBEKLRSKMKMKWATFARTGNPNGN 510
DB 512 TKFAFYVADHAEVRFVFGAFLKGDIVMEGATEEBEKLRSKMKMKWATFARTGNPNGN 571
QY 511 DLSLWPAVNTTEOYLQDLNMSIGORLKEPRVDFWSTTPIILISASDMHSPLSITFLS 570
DB 572 DLSLWPAVNTTEOYLQDLNMSIGORLKEPRVDFWSTTPIILISASDMHSPLSITFLS 631

```

```

QY 571 LLOPFFFCAP 581
DB 632 LLOPFFFCAP 642

```

```

RESULT 6
US-10-451-168-92
/ Sequence 92, Application US/10451168
/ Publication No. US20040091969A1
/ GENERAL INFORMATION:
/ APPLICANT: SMITHKLINE BEECHAM CORPORATION
/ APPLICANT: GLAXO GROUP LIMITED
/ TITLE OF INVENTION: NOVEL COMPOUNDS
/ FILE REFERENCE: GP50039
/ CURRENT APPLICATION NUMBER: US/10/451,168
/ PRIOR FILING DATE: 2003-11-12
/ PRIOR APPLICATION NUMBER: PCT/US01/49232
/ PRIOR FILING DATE: 2000-12-17
/ PRIOR APPLICATION NUMBER: 60/256,710
/ PRIOR FILING DATE: 2000-12-19
/ PRIOR APPLICATION NUMBER: 60/257,048
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 60/260,482
/ PRIOR FILING DATE: 2001-01-09
/ PRIOR APPLICATION NUMBER: 60/264,922
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/266,797
/ PRIOR FILING DATE: 2001-02-06
/ PRIOR APPLICATION NUMBER: 60/276,988
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/281,535
/ PRIOR FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: 60/289,622
/ PRIOR FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 92
/ LENGTH: 581
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-451-168-92

```

```

Query Match      84.7%; Score 2607.5; DB 15; Length 581;
Best Local Similarity 91.4%; Pred. No. 1e-233;
Matches 499; Conservative 7; Mismatches 25; Indels 15; Gaps 3;

QY 12 WCEFLILQPLGHRQWKTGPSAEGPQPNTRLGWTIGKQVTVLGSPPVNVFLGPPAP 71
DB 14 WAIWVLAAP-----TKGPSAEGPQPNTRLGWTIGKQVTVLGSPPVNVFLGPPAP 65
QY 72 PLGSIRFTNPQAPSPMDNIREATSYPNLCLONSEWMLLDQMLKVHYPKFGVSEDCLYLN 131
DB 66 PLGSIRFTNPQAPSPMDNIREATSYPNLCLONSEWMLLDQMLKVHYPKFGVSEDCLYLN 125
QY 132 IYAPAHADTGSKLPVYWFPGGAFTKGSASIFDGSALAAYEDVLVVVQYRLGIFGFTT 191
DB 126 IYAPAHADTGSKLPVYWFPGGAFTKGSASIFDGSALAAYEDVLVVVQYRLGIFGFTT 185
QY 192 WVOQNIIEFFGDDPSVYTFGFSAGAISSVSLILSPMAKGLF 251
DB 186 WVOQNIIEFFGDDPSVYTFGFSAGAISSVSLILSPMAKGLF 245
QY 251 HKAIMESGVALIPLYBAHDYKSEDIQVVAHFPGNNASDSEALLRCLRTKPSKELTLISQ 311
DB 246 HKAIMESGVALIPLYBAHDYKSEDIQVVAHFPGNNASDSEALLRCLRTKPSKELTLISQ 305
QY 312 KTKSFTFRVVDGAFPPNEPLDLSQKAFKAIPIISIIIGVNNHCEGFLIPMKCAPETILSSGN 369
DB 306 KTKSFTFRVVDGAFPPNEPLDLSQKAFKAIPIISIIIGVNNHCEGFLIPMKCAPETILSSGN 365
QY 369 KSMAL-----HLIQLIHLIPQYTLHVAANEYFHDKSLTEIRDSLDDLDGDFVFPVVALI 424

```

Db	366	RDALAASTAGHFRHROHLPQYLMHVAAREFYHDKSLTEIRDSLLDLDGDFEVVPAL	425
Qy	425	TARYHRDGAAPYFYEFHRHPOCFEDTKPAFKADHAEVRFYGGAFLKGDIVMEGAT	484
Db	426	TARYHRDGAAPYFYEFHRHPOCFEDTKPAFKADHAEVRFYGGAFLKGDIVMEGAT	485
Qy	485	EEBKLSLKKKKMYATPARTGPNPNDLSLMPAYNLTEOYLQDLNMSLGORLKEBRDV	544
Db	486	EEBKLSLKKKKMYATPARTGPNPNDLSLMPAYNLTEOYLQDLNMSLGORLKEBRDV	545
Qy	545	WTSTRIP 550	
Db	546	WVTGTP 551	

RESULT 7
US-10-094-749-2375

```

Sequence 2375, Application US/10094749
Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAOBU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 2375
LENGTH: 525
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2375

```

Query Match	84.6%;	Score 2606;	DB 15;	Length 525;
Best Local Similarity	90.9%;	Pred. No. 1.2e-233;		
Matches 501; Conservative	0;	Mismatches 0;	Indels 50;	Gaps 1

Qy	3	GPSAEGPQPNRLGMIQOKQVTVLSPVYVNYFLGVFPAPAPPLGSLRFTNPPQASPMNDL	90
Db	25	GPSAEGPQPNRLGMIQOKQVTVLSPVYVNYFLGVFPAPAPPLGSLRFTNPPQASPMNDL	84
Qy	91	REATSYPNLCIONSEWMLLDQHLKHYHYPKFGVSEDCLYNITYAPAHADTGS.KYLV.VMF	150
Db	85	REATSYPNLCIONSEWMLLDQHLKHYHYPKFGVSEDCLYNITYAPAHADTGS.KYLV.VMF	144
Qy	151	PGGAKTTSASIFDGSALAAVEDLVVVVQYQYRLGIFGEFTTMDQAH.PGNMA.FDQVAAALS	210
Db	145	PGGAKTTSASIFDGSALAAVEDLVVVQYQYRLGIFGEFTTMDQAH.PGNMA.FDQVAAALS	204
Qy	211	VWQXNIIEFFGQDPSSVTIIFGSAGAISSVSLILSPMAKGLFHKAINEGCVALLPYLEAHD	270

```

Db      205  WVKQNIIEFFGGDPSBVTITGEBAGAIWVSLSLSPAKLFFKAINESGVAIIPIYLEAH 264
Qy      271  YKESDLOVVAHPCGNMNSDSEBALLRCITRKPSKELLJLSQTKTSFTRVVDGAFPPNEPL 330
Db      265  YKESDLOVVAHPCGNMNSDSEBALLRCITRKPSKELLJLSQTKTSFTRVVDGAFPPNEPL 334
Qy      331  DLSGAKAKAPSTITGVNNHECGPILPMKEAPEILSGNSKSLALHILIONLHPPQYLH 390
Db      325  DLSGAKAKAPSTITGVNNHECGPILPMKEAPEILSGNSKSLALHILIONLHPPQYLH 384
Qy      391  VANEYFHDKHSITETIRDSILDLGDPVFVPPALITARYRDAGAPVYFYEFRRHPQCFED 450
Db      385  VANEYFHDKHSITETIRDSILDLGDPVFVPPALITARYR----- 424
Qy      451  TKPAFVKADHADVRFVFGAFLKGDIVWFEGATEBEKLSRMMKYMATFARTGNPNCN 510
Db      425  -----ECATEBEKLSRMMKYMATFARTGNPNCN 454
Qy      511  DLSLMPAYNLTEQYIQLDLNMSLGQRLKEPRVDFTSTPLILSASDMLHSPSLSTFLS 570
Db      455  DLSLMPAYNLTEQYIQLDLNMSLGQRLKEPRVDFTSTPLILSASDMLHSPSLSTFLS 514
Qy      571  LLQPFEEFCAP 581
Db      515  LLQPFEEFCAP 525

```

```

RESULT 8
US-10-114-270-196
; Sequence 196, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:

```

1 APPLICANT: Guo, Xiaojia
 2 APPLICANT: Kekuda, Ramesh
 3 APPLICANT: Miller, Charles E.
 4 APPLICANT: Malyankar, Uriel M.
 5 APPLICANT: Spytek, Kimberly A.
 6 APPLICANT: Paturajan, Meera
 7 APPLICANT: Liu, Ziaohong
 8 APPLICANT: Gusev, Vladimir Y.
 9 APPLICANT: Li, Li
 10 APPLICANT: Vernet, Corine
 11 APPLICANT: Zerhusen, Bryan D.
 12 APPLICANT: Gorman, Linda
 13 APPLICANT: Shenoy, Suresh G.
 14 APPLICANT: Pena, Carol E.A.
 15 APPLICANT: Smithson, Glenda
 16 APPLICANT: Burgess, Catherine E.
 17 APPLICANT: Gerlach, Valerie
 18 APPLICANT: Padigaru, Muralidhara
 19 APPLICANT: Shimkets, Richard A.
 20 APPLICANT: Gangolli, Esba A.
 21 APPLICANT: Taupler Jr., Raymond J.
 22 APPLICANT: Caeman, Stacie J.
 23 APPLICANT: Ji, Weizhen
 24 APPLICANT: Anderson, David W.
 25 APPLICANT: Liete, Mario W.
 26 APPLICANT: Rastelli, Luca
 27 APPLICANT: Edinger, Shlomit R.
 28 APPLICANT: Stone, David J.
 29 APPLICANT: MacDougall, John R.
 30 APPLICANT: Rothenberg, Mark E.
 31 TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same
 32 FILE REFERENCE: 22402-329C
 33 CURRENT APPLICATION NUMBER: US/10/114,270
 34 CURRENT FILING DATE: 2002-11-27
 35 PRIOR APPLICATION NUMBER: 60/281,086
 36 PRIOR FILING DATE: 2001-04-03
 37 PRIOR APPLICATION NUMBER: 60/281,136
 38 PRIOR FILING DATE: 2001-04-03
 39 PRIOR APPLICATION NUMBER: 60/281,863
 40 PRIOR FILING DATE: 2001-04-05
 41 PRIOR APPLICATION NUMBER: 60/281,906

```

; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 196
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-196

```

```

Query Match      84.4%; Score 2597.5; DB 15; Length 581;
Best Local Similarity 94.3%; Pred. No. 8.7e-233;
Matches 497; Conservative 4; Mismatches 17; Indels 9; Gaps 3;

```

```

QY 31 GGSAGSPQNTNLGWTGKQVTVLGSPPVNVFLGVFAAPLGSIRFTNPQPASPMNDL 90
DB 27 GGSAGSPQNTNLGWTGKQVTVLGSPPVNVFLGVFAAPLGSIRFTNPQPASPMNDL 86
QY 91 REATSYPNLCQNSEWLLLDQMLKVHYPKFVSEDCLYNTIYAPAHADTGSKLPVLVWF 150
DB 87 REATSYPNLCQNSEWLLLDQMLKVHYPKFVSEDCLYNTIYAPAHADTGSKLPVLVWF 146
QY 151 PGAFKTSASIFDSALAAVEDLVVVVQYRLGIFGFTTWDQAPGNMAFKDOVALS 210
DB 147 PGAFKTSASIFDSALAAVEDLVVVVQYRLGIFGFTTWDQAPGNMAFKDOVALS 206
QY 211 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAND 270
DB 207 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAND 266
QY 271 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKKELLTLOSQTSPFRVNDGAPFPNEPL 330
DB 267 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKKELLTLOSQTSPFRVNDGAPFPNEPL 326
QY 331 DLSQKAFKAIPSIIGVNNHECGFLPMKE--APEILSGSNKSLAL-----HLIQNLIHI 383
DB 327 DLSQKAFKAIPSIIGVNNHECGFLPMKE--APEILSGSNKSLAL-----HLIQNLIHI 384
QY 384 PPOYLHVANEYFHDKSLTEIRSLDLGDFVFPVVALITARYHRDAGAPVYEFRRH 443
DB 385 PPOYLHVANEYFHDKSLTEIRSLDLGDFVFPVVALITARYHRDAGAPVYEFRRH 444
QY 444 RPOCFEDTKPAFVADNADEVRFVFGAFLKGDIVMEGATEEBEKLSTRKMKMTAFAR 503
DB 445 RPOCFEDTKPAFVADNADEVRFVFGAFLKGDIVMEGATEEBEKLSTRKMKMTAFAR 504
QY 504 TGNNGNDLSIMPRVYNLTQYLDLQNLMSLQGLKEPRVDFWTSTIP 550
DB 505 TGNNGNDLSIMPRVYNLTQYLDLQNLMSLQGLKEPRVDFWTSTIP 551

```

RESULT 9

```

US-10-381-898-2
; Sequence 2, Application US/10381898
; Publication No. US20040086887A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKY, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
; APPLICANT: HAPALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Fatrah A.; IAL, Preeti G.;

```

```

; APPLICANT: LEE, Ernestine A.; LU, Dying Aina M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
; APPLICANT: SANJANWALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOLEV, Catherine M.;
; APPLICANT: CHAWLA, Narinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: YANG, Junning; YAO, Monique; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: US/10/381,898
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/236,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473645CD1
US-10-381-898-2

```

```

Query Match      84.1%; Score 2588.5; DB 15; Length 618;
Best Local Similarity 94.4%; Pred. No. 6.6e-232;
Matches 491; Conservative 1; Mismatches 5; Indels 23; Gaps 1;

```

```

QY 31 GGSAGSPQNTNLGWTGKQVTVLGSPPVNVFLGVFAAPLGSIRFTNPQPASPMNDL 90
DB 27 GGSAGSPQNTNLGWTGKQVTVLGSPPVNVFLGVFAAPLGSIRFTNPQPASPMNDL 151
QY 91 REATSYPNLCQNSEWLLLDQMLKVHYPKFVSEDCLYNTIYAPAHADTGSKLPVLVWF 150
DB 87 REATSYPNLCQNSEWLLLDQMLKVHYPKFVSEDCLYNTIYAPAHADTGSKLPVLVWF 151
QY 151 PGAFKTSASIFDSALAAVEDLVVVVQYRLGIFGFTTWDQAPGNMAFKDOVALS 210
DB 147 PGAFKTSASIFDSALAAVEDLVVVVQYRLGIFGFTTWDQAPGNMAFKDOVALS 271
QY 211 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAND 270
DB 207 WQKNIEFFGDPSSVTTFGESAGAISSVSLTSPMAKGLFHKAIMESGVAIIPYLEAND 331
QY 271 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKKELLTLOSQTSPFRVNDGAPFPNEPL 330
DB 267 YEKSEDLQVNAHFCNNASDSEALRLCRLTPSKKELLTLOSQTSPFRVNDGAPFPNEPL 391
QY 331 DLSQKAFKAIPSIIGVNNHECGFLPMKEAPEILSGSNKSLALHLIQNLIHI 390
DB 327 DLSQKAFKAIPSIIGVNNHECGFLPMKEAPEILSGSNKSLALHLIQNLIHI 428
QY 384 PPOYLHVANEYFHDKSLTEIRSLDLGDFVFPVVALITARYHRDAGAPVYEFRRH 450
DB 385 PPOYLHVANEYFHDKSLTEIRSLDLGDFVFPVVALITARYHRDAGAPVYEFRRH 488
QY 444 RPOCFEDTKPAFVADNADEVRFVFGAFLKGDIVMEGATEEBEKLSTRKMKMTAFAR 510
DB 445 RPOCFEDTKPAFVADNADEVRFVFGAFLKGDIVMEGATEEBEKLSTRKMKMTAFAR 510

```

Db 489 TKPAFVKADHADEVFVEFGAFGLKGDIVMFEAGATEEKKLSRKMKRYWATFARTGNPN 548
 Qy 511 DLSMPAYNLTEOYLQDLNMSLGORLKEPRVDFTSTIP 550
 Db 549 DLSMPAYNLTEOYLQDLNMSLGORLKEPRVDFTSTIP 588

RESULT 10

US-10-104-047-2219
 ; Sequence 2219, Application US/10104047
 ; Publication No. US20030236392A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20030236392A1e1 full length cDNA
 ; FILE REFERENCE: H1-A0105
 ; CURRENT FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2219
 ; LENGTH: 469
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-104-047-2219

Query Match 79.7%; Score 2454; DB 15; Length 469;
 Best Local Similarity 99.6%; Pred. No. 1.5e-219;
 Matches 467; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 113 MLKHYPRKGVSEDCLYNTIYAPAHADTGSKLPLYVWFGAFKTSASIFDGSALAYE 172
 Db 1 MLKHYPRKGVSEDCLYNTIYAPAHADTGSKLPLYVWFGAFKTSASIFDGSALAYE 60
 Qy 173 DVLVVVVVYRIGIFGFTTMDQAHAPGNWAFKQVVAALSMVQKNIFFGSDPSVTIIFGSS 232
 Db 61 DVLVVVVVYRIGIFGFTTMDQAHAPGNWAFKQVVAALSMVQKNIFFGSDPSVTIIFGSS 120
 Qy 233 AGAISVSSLLISPMKAGLPHKAIMESGVAIIPYLEAHDIKESEDLQVVAHFCGNASDSE 292
 Db 121 AGAISVSSLLISPMKAGLPHKAIMESGVAIIPYLEAHDIKESEDLQVVAHFCGNASDSE 180
 Qy 293 ALALCLRTKPSKELLTTSOKTSSFTRVVDGAFPEPNEPLDLSOKAFKAIPTIIGVNNHC 352
 Db 181 ALALCLRTKPSKELLTTSOKTSSFTRVVDGAFPEPNEPLDLSOKAFKAIPTIIGVNNHC 240
 Qy 353 GFLPMPKAPRLLSGSNKSLALHLIQLIHTIPQYLHLVANEYFADKSLTEIRDSLLDL 412
 Db 241 GFLPMPKAPRLLSGSNKSLALHLIQLIHTIPQYLHLVANEYFADKSLTEIRDSLLDL 300
 Qy 413 LGDVFVVPALITARYHRDAGAPVYFEFRHRPQCFEDTKPAFVADHADEVFVFGAF 472
 Db 301 LGDVFVVPALITARYHRDAGAPVYFEFRHRPQCFEDTKPAFVADHADEVFVFGAF 360
 Qy 473 LKGDIVMEEGATEEKKLSRKMKRYWATFARTGNNGNDLSMPAYNLTEOYLQDLNMS 532
 Db 361 LKGDIVMEEGATEEKKLSRKMKRYWATFARTGNNGNDLSMPAYNLTEOYLQDLNMS 420
 Qy 533 LGORLKEPRVDFTSTIPILISASDMLHSPLSLTFLSLQFPFFFCAP 581
 Db 421 LGORLKEPRVDFTSTIPILISASDMLHSPLSLTFLSLQFPFFFCAP 469

RESULT 11
 US-10-233-933A-2
 ; Sequence 2, Application US/10233933A
 ; Publication No. US20040214171A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamashita, Tetsuro
 ; APPLICANT: Miyazaki, Masao
 ; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
 ; FILE REFERENCE: SHIG FP02US006
 ; US-10-233-933A-2

; CURRENT APPLICATION NUMBER: US/10/233,933A
 ; CURRENT FILING DATE: 2002-09-03
 ; PRIOR APPLICATION NUMBER: JP2002-057908
 ; PRIOR FILING DATE: 2002-04-03
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 542
 ; TYPE: PRT
 ; ORGANISM: Felis catus
 ; US-10-233-933A-2

Query Match 64.5%; Score 1985; DB 16; Length 542;
 Best Local Similarity 71.7%; Pred. No. 1e-175;
 Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;

Qy 31 GPSAEGPQNRTRLGWITQSKQVTVLGSPPVNVFLGVPPAAPPLGSLRTNPPAPSPMDNL 90
 Db 22 GPADAPVRSSTRLGVRNRKQTTVLGSTVPMVFLGI PYAAPPLGSLRFRQPPALPGNDP 81
 Qy 91 REATSVPLCLQNSFWLLDDHMLKVHPKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 150
 Db 82 RNATSVPLCLQNSFWLLDDHMLKVHPKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 141
 Qy 151 PGAFKTSASIFDGSALAAVEDVLLVVOYRIGIFGFTTMDQAHAPGNWAFKQVVAALS 210
 Db 142 PGAFKTSASIFDGSALAAVEDVLLVVOYRIGIFGFTTMDQAHAPGNWAFKQVVAALS 201
 Qy 211 WVQKNIEFFGSDPSVTIIFGSSAGAI SVSLLISPMKAGLPHKAIMESGVAIIPYL-ENH 269
 Db 202 WVRDNIIEFFGSDPSVTIIFGSSAGAI SVSLLISPMKAGLPHKAIMESGVAIIPYL-ENH 261
 Qy 270 DYKSEDLQVVAHFCGNASDSEALRLCTKPSKELLTTSOKTSSFTRVVDGAFPEPNEP 329
 Db 262 DYKSEDLQVVAHFCGNASDSEALRLCTKPSKELLTTSOKTSSFTRVVDGAFPEPNEP 321
 Qy 330 LDLSQKAFKAIPTIIGVNNHCGLPMPKAPRLLSGSNKSLALHLIQLIHTIPQYLH 389
 Db 322 VALLTQKAFNSVPSIIGVNNHCAPL-STERSLISGSNLSALYVHTPLNIFQYLH 380
 Qy 390 LVANEYFADKSLTEIRDSLLDLGDFVFPVVPALITARYHRDAGAPVYFEFRHRPQCFE 449
 Db 381 LVANEYFADKSLTEIRDSLLDLGDFVFPVVPALITARYHRDAGAPVYFEFRHRPQCFE 440
 Qy 450 DTKPAFVADHADEVFVFGAFGLKGDIVMFEAGATEEKKLSRKMKRYWATFARTGNPN 509
 Db 441 DTKPAFVADHADEVFVFGAFGLKGDIVMFEAGATEEKKLSRKMKRYWATFARTGNPN 500
 Qy 510 NDLSMPAYNLTEOYLQDLNMSLGORLKEPRVDFTSTI 549
 Db 501 NDLSMPAYNLTEOYLQDLNMSLGORLKEPRVDFTSTI 540

RESULT 12
 US-10-233-933A-4
 ; Sequence 4, Application US/10233933A
 ; Publication No. US20040214171A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamashita, Tetsuro
 ; APPLICANT: Miyazaki, Masao
 ; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
 ; FILE REFERENCE: SHIG FP02US006
 ; CURRENT FILING DATE: 2002-09-03
 ; PRIOR APPLICATION NUMBER: JP2002-057908
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 542
 ; TYPE: PRT
 ; ORGANISM: Felis catus
 ; US-10-233-933A-4

```

Query Match      64.5%; Score 1985; DB 16; Length 542;
Best Local Similarity 71.7%; Pred. No. 1e-175;
Matches 373; Conservative 60; Mismatches 85; Indels 2; Gaps 2;

QY 31 GPSABGPQNTRLGMIQKQVTVLGSPPVNVFLGVFPFAAPPLGSLRTNPOPASFWDL 90
DB 22 GPAAADAPVSTRIGWRGKQTVLGSSTVPVNMFLGIPYAAPPLGRLFPQPKALPGNDF 81
QY 91 REATSPNCLONSEWLLDDQHLKXHYPRFGVSEDCVLYNTIYAPHAATGSLPLVWVF 150
DB 82 RNATSPKLCFQDLEWLVSIQHLKXRYPLENSEDCVLYNTIYAPHAADNGSLPLVWVF 141
QY 151 PGCAFRTGSASIFDGSALAAVEDLVVVVQYRLGIPGFPTWDOHAPGMMAFDDOVAALS 210
DB 142 PGCAFRTGSASIFDGSALAAVEDLVVVQYRLGIPGFPTDDEHAKGMALLDOVAALT 201
QY 211 WQKNIPEFGDPSYTTITGESAGATSVSLLISPMAKGLFKKAIMESGVAIPYL-EAH 269
DB 202 WVRDNIPEFGDPSYTTITGESAGATSVSLLISPIANGLFKKAIMESGVALPLLMRPP 261
QY 270 DYKESDLOVNAHFCGNNASDSEALLRCLRTKPSKELLTSQKTSFTRVVVGAFPPNRP 329
DB 262 GBERKDDLOVNAHFCGNNASDSEALLRCLRTKPSKELLTSQKTSFTRVVVGAFPPNRP 321
QY 330 LDLSQKAFKALPSIIGVNNHCGFLLPWKEAPEILSGSNKSLALHILIONLIHIPQYLAH 389
DB 322 VALTQKAFNSVPSIIGVNNHCGFLLPWKEAPEILSGSNKSLALHILIONLIHIPQYLAH 380
QY 390 LVANEFYHDKSLTEIRDSLLDLGDVFPVVPALITARYHRDAGAPVTFYEFRRPQCFE 449
DB 381 LVADHYFYNKHSFVIRDSFLDLGDVFPVVPALITARYHRDAGAPVTFYEFRRPQCFE 440
QY 450 DTKPAFVADHAEVRFVFGAFLKGDIVMEFEGATEEBSKLSRKMMKWTAPFARTGNNG 509
DB 441 DTRPAFVADHAEVRFVFGAFLKGDIVMEFEGATEEBSKLSRKMMKWTAPFARTGNNG 500
QY 510 NDLSIWPAYNLTQOYLQDLNNSLQGLKEPRVDFWSTI 549
DB 501 EGVPLMPAYTQSEQYTKIDLVSVGKLEQVEFEWMMNTI 540

RESULT 13
US-10-451-168-93
; Sequence 93, Application US/10451168
; Publication No. US20040091969a1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; PRIOR FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-451-168-93

Query Match      57.8%; Score 1780; DB 15; Length 356;
Best Local Similarity 100.0%; Pred. No. 7.1e-157;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ILSPAKGLFKKAIMESGVAIIPYLEAHDYKESDLOVNAHFCGNNASDSEALLRCLRTK 301
DB 17 ILSPAKGLFKKAIMESGVAIIPYLEAHDYKESDLOVNAHFCGNNASDSEALLRCLRTK 76
QY 302 PSKELLTSQKTSFTRVVVGAFPPNEPLDLSQKAFKALPSIIGVNNHCGFLLPWKEA 361
DB 77 PSKELLTSQKTSFTRVVVGAFPPNEPLDLSQKAFKALPSIIGVNNHCGFLLPWKEA 136
QY 362 PEILSGSNKSLALHILIONLIHIPQYLAHLYANEFYHDKSLTEIRDSLLDLGDVFPVP 421
DB 137 PEILSGSNKSLALHILIONLIHIPQYLAHLYANEFYHDKSLTEIRDSLLDLGDVFPVP 196
QY 422 ALITARYHRDAGAPVTFYEFRRPQCFEDTKPAFVKADHAEVRFVFGAFLKGDIVMEF 481
DB 197 ALITARYHRDAGAPVTFYEFRRPQCFEDTKPAFVKADHAEVRFVFGAFLKGDIVMEF 256
QY 482 GATEEBSKLSRKMMKWTAPFARTGNPNNDLSIMPAYNLTQOYLQDLNNSLQGLKEPR 541
DB 257 GATEEBSKLSRKMMKWTAPFARTGNPNNDLSIMPAYNLTQOYLQDLNNSLQGLKEPR 316
QY 542 VDFWSTIIPILSASDMLHSPLSTPLSLQFPFFFCAP 581
DB 317 VDFWSTIIPILSASDMLHSPLSTPLSLQFPFFFCAP 356

RESULT 14
US-09-925-298-689
; Sequence 689, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 689
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-689

Query Match      39.9%; Score 1230; DB 9; Length 549;
Best Local Similarity 46.7%; Pred. No. 2.6e-105;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

QY 31 GPSABGPQNTRLGMIQKQVTVLGSPPVNVFLGVFPFAAPPLGSLRTNPOPASFWDL 90
DB 16 GODSASPIRTHTGVLGSLVHVKGANAGVOTFLGIPPAKPEPLGLRPAPEPPSPSGV 75

```

```

Qy 91 REATSYVNLICLN-----SEWILLDOHMLKVHYPKRGVSEDCLYINIVAPAHADGSKLP 145
Db 76 RDGITHPAMCLODLTVAVESEFL-----SQNNMTFPEDSMSSEDCLYISTYTPAHSHGSLNP 131
Qy 146 VLVWFPFGAFKTKGASIFDGSALAAVEDVLYVVVQYRIGIFGFPTTMOHAGNAPFQDQ 205
Db 132 VVVVTHGGALVFGMASTLYDGSMLALENVVVVYIYRIGVGLGFSTGDKHATGNMGYLDQ 191
Qy 206 VVALSMVQKNIEFFGDPSSVTITFGESAGATSVSSLLSPMAKGLFHKAIMSGVAILIPY 265
Db 192 VVALSMVQKNIAHFGNDPRVTITFGESAGATSVSSLLSPISQGLFHGAIMSGVAILIPG 251
Qy 266 LEADHYKSEDL--OVVAHFCGNNASDSEBALRCURTKPSKELLTLTQKTSFTRVVDA 323
Db 252 LIA-----SSADVI STVVANLSACDQVDSBALVGCIRGSKKEIILAINRPFKMI PGVVG 307
Qy 324 FFPNEPDLLOSOKAKAIPSIIGVNNHCGFLP---MKEAPETLSSGNSKLALHLION 379
Db 308 FLPRHPQELIASADQVPVPSIVGVNNNEFGWLI PKVMRIYDTQKEMDRASQALQKMLT 367
Qy 380 ILHIPPOYLHIVANEYFHDKISLTIIRDLSLDLDGDFVFPVVPALITARYHRDAGAPVYFY 439
Db 368 LLMPTTGDLIREYIGDNGDPQTLQAQFQEMMADSMFVLPALQVAHF-QCSRAPVYFY 426
Qy 440 EFRHRPOCFEDTKPAFVADHADEVRFVFGAFLKGDIVMEGATEEBSKLSRKMMKTYA 499
Db 427 EFQHPQSWLKNIRPPhMKADHDELPLVYF-RSFGGNVYKF---TEEBEQLSRKMMKTYA 482
Qy 500 TPARGTGNVNDLSLMPAYNLTEQYLQDLNMSLQRLKEPRVDWTSTIP 550
Db 483 NFARNGNPNBGGLPMPPLFDQEEQYLQNLQPAVGRALKAHRLQFWKXALP 533

```

RESULT 15
US-10-102-806-689

```

; Sequence 689, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103PICI
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 689
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-689

```

Query Match 39.9%; Score 1230; DB 14; Length 549;

Best Local Similarity 46.7%; Pred. No. 2,6e-105;

Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

```

Qy 31 GPSAAGPQANTLGMVIGKQVTVLGSFVNVVFLVGPAPPLGSLRFTNPOPASPMWNL 90
Db 16 GQDSASPIRTHYTGVLGSLVHYKGNAGNVQTFGLGIPAKPPLGRLPAPPEPESWGSV 75
Qy 91 REATSYVNLICLN-----SEWILLDOHMLKVHYPKRGVSEDCLYINIVAPAHADGSKLP 145

```

```

Db 76 RDGITHPAMCLODLTVAVESEFL-----SQNNMTFPEDSMSSEDCLYISTYTPAHSHGSLNP 131
Qy 146 VLVWFPFGAFKTKGASIFDGSALAAVEDVLYVVVQYRIGIFGFPTTMOHAGNAPFQDQ 205
Db 132 VVVVTHGGALVFGMASTLYDGSMLALENVVVVYIYRIGVGLGFSTGDKHATGNMGYLDQ 191
Qy 206 VVALSMVQKNIEFFGDPSSVTITFGESAGATSVSSLLSPMAKGLFHKAIMSGVAILIPY 265
Db 192 VVALSMVQKNIAHFGNDPRVTITFGESAGATSVSSLLSPISQGLFHGAIMSGVAILIPG 251
Qy 266 LEADHYKSEDL--OVVAHFCGNNASDSEBALRCURTKPSKELLTLTQKTSFTRVVDA 323
Db 252 LIA-----SSADVI STVVANLSACDQVDSBALVGCIRGSKKEIILAINRPFKMI PGVVG 307
Qy 324 FFPNEPDLLOSOKAKAIPSIIGVNNHCGFLP---MKEAPETLSSGNSKLALHLION 379
Db 308 FLPRHPQELIASADQVPVPSIVGVNNNEFGWLI PKVMRIYDTQKEMDRASQALQKMLT 367
Qy 380 ILHIPPOYLHIVANEYFHDKISLTIIRDLSLDLDGDFVFPVVPALITARYHRDAGAPVYFY 439
Db 368 LLMPTTGDLIREYIGDNGDPQTLQAQFQEMMADSMFVLPALQVAHF-QCSRAPVYFY 426
Qy 440 EFRHRPOCFEDTKPAFVADHADEVRFVFGAFLKGDIVMEGATEEBSKLSRKMMKTYA 499
Db 427 EFQHPQSWLKNIRPPhMKADHDELPLVYF-RSFGGNVYKF---TEEBEQLSRKMMKTYA 482
Qy 500 TPARGTGNVNDLSLMPAYNLTEQYLQDLNMSLQRLKEPRVDWTSTIP 550
Db 483 NFARNGNPNBGGLPMPPLFDQEEQYLQNLQPAVGRALKAHRLQFWKXALP 533

```

Search completed: June 15, 2005, 09:46:09
Job time : 88 secs

This Page Blank (uspto)

PF 18-DEC-2001; 2001MO-US049075.
XX
PR 18-DEC-2000; 2000US-0256369P.
PR 28-MAR-2001; 2001US-0279508P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Siles-Santiago I;
XX
DR MPI; 2002-547936/58.
DR P-PSDB; ABB93537.
XX

53010 nucleic acids, useful for diagnosing and treating e.g. vascular diseases, autoimmune diseases, or neurodegenerative diseases, as surrogate markers, in tissue typing and chromosome mapping.

Claim 1; Page 109-110; 122pp; English.

The present nucleotide sequence, the coding region of which is also claimed, is that of cDNA encoding human 53010, a novel member of the carboxylesterase family. The invention provides 53010 nucleic acids, antisense molecules, expression vectors, host cells, transgenic animals, 53010 proteins, fusion proteins, antigenic peptides, anti-53010 antibodies and methods for detecting the presence of 53010 polypeptides or nucleic acids, of identifying a compound that binds to the 53010 polypeptide, and of modulating the activity of the polypeptide. The 53010 nucleic acids and polypeptides can act as novel diagnostic and therapeutic agents for controlling disorders involving aberrant or deficient hydrolysis of carboxylic esters. As 53010 mRNA is highly expressed in the central and peripheral nervous system, and its expression is regulated in some rodent pain models, 53010 molecules can also act as novel diagnostic targets and therapeutic agents for controlling neurological disorders, such as pain-related disorders. A claimed method of treating or preventing a disorder (especially a pain-related disorder) characterised by aberrant activity of a 53010-expressing cell involves administering a compound that modulates 53010 activity or expression. 53010 nucleic acids are also useful in chromosome mapping, tissue typing, in forensic biology, prognostic assays, in arrays, for detection of variations or mutations, as surrogate markers and in pharmacogenomics

Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,71e-282 Length: 2158
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-023-515-2 (1-581) x ABBN84302 (1-2158)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCySPhePheLeuIleLeuGlnPro 20
Db 96 ATGCCACAGGAGCTTACTTCATCTGCTTCAACAAATGGTCTTTTCTCGATTCTCCAGCC 155
QY 21 LeuLeuGlnIleSarGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnAlaAsn 40
Db 156 CTGTTGGGACACACAGCTGGGGAAGAACTGGGCTTCTGCTGAAGGCCACAGAGAAC 215
QY 41 ThrArgLeuGlyTrpIleGlnGlyLeuGlnValThrValLeuGlySerProValProVal 60
Db 216 ACCAGGCTGGGATGATTCAAGGCAAGCAAGTCACTGTGGGAAGCCCTGTGCTGTG 275
QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerIleuAlaGpPheThrAsn 80
Db 276 AACCTGTTCTCTGGAGTCCTCTTGCTGCTCCCGCTGGGATCCCTGGCAATTTAACAAAC 335
QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnAlaThrSerTrpProAsnLeuCyS 100
Db 336 CCGCAGCTGTGATGCCCTGGGATTAATTGGAGAAAGCACTCTACCTTAATTGTGTC 395

QY 101 LeuGlnaenSerGlnTrpLeuLeuLeuAspGlnIleSweLeuLeuValIleTrpProLyS 120
Db 336 CTCAGAACTCAGAGTGGCTGCTTATGATCAACACATGCTCAAGTGCATTAACCCGAAA 455
QY 121 PheGlyValSerGlyAspCySPheLeuLeuLeuIleTrpAlaProAlaIleAlaAspThr 140
Db 456 TTCCGAGTGTCAAGAACTGCTCTTCACTTGAACATCTAGCCCTCCACGCCGATACA 515
QY 141 GlySerIleLeuProValLeuValTrpPheProGlyGlyAlaPheLeuThrGlySerAla 160
Db 516 GGCTCCAAAGTCCCGCTGTGTGTGTGTCCACGAGAGTGCCTTCAAGACTGGCTGACC 575
QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValValGln 180
Db 576 TCCATCTTGAATGGATCCCGCTGCTGCTCATATAGACGCTGCTGTGTGTGTGTGTCAG 635
QY 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnIleAlaProGlyAsnTrp 200
Db 636 TACCGGCTAGGAATTTTGGTTTCTTCCACCAATGGATCAGCATGCTCCGAGGAATCG 695
QY 201 AlaPheIleAspGlnValAlaAlaLeuSerTrpValGlnIleAsnIleGluPheGly 220
Db 696 GCCTTCAAGGACCAAGTGGCTGCTGTCCCTGGGTCCAGAAACATCGAGTTCTTCGT 755
QY 221 GlyAspProSerSerValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSer 240
Db 756 GGGACCCCAAGCTTGTGACCATCTTGGGAGTCCCGGAGCCATTAAGTGTTCATG 815
QY 241 LeuIleLeuSerProMetAlaIleGlyLeuPheIleIleValAlaIleMetGlySerGlyVal 260
Db 816 CTATACGTGCTCCCATGAGCAAAAGCTTATTCACAAACCATCAGAGAGTGGGTG 875
QY 261 AlaIleIleProTrpTrpLeuGlnAlaIleAspTrpGlyIleSerGlyLeuGlnValVal 280
Db 876 GCCATCATCCCTTACCTGAGGCCCATGATTAAGAAAGTGAAGGACCTGCAAGTGTG 935
QY 281 AlaIlePheCySPheValaenAlaSerAspSerGlyAlaLeuLeuArgCySPheArgTrp 300
Db 936 GCACATTTCTGTGTGAACATGCGTCAAGCTCTGAGGCCCTTGTGAGTCTTGAAGACA 995
QY 301 LysProSerIleGlyLeuLeuTrpLeuSerGlnIleTrpLysSerPheThrArgValVal 320
Db 996 AAACCTCCAAAGGAGCTGACCTCAGCCACAGAAACAAAGCTTTCACTCAGAGTGTG 1055
QY 321 AspGlyValaPhePheProAsnGlnProLeuAspLeuLeuSerGlnIleValaPheVala 340
Db 1056 GATGGTGTCTTCTTCTTCAATGAGCCCTAGATCTATGTCTCAGAAAGCATTTAAAGCA 1115
QY 341 IleProSerIleIleGlyValaenAlaenIleGlyCySPheLeuLeuProMetLysGly 360
Db 1116 ATTCTTCCATCATCGAGGTCATTAACCAAGATGTGGCTTCTGCTGCTTATGAAGAG 1175
QY 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuIleGlnaenIle 380
Db 1176 GCTCTGAGATCTCTCAGTGGCTCAACAAAGTCCCTTGCCCTCATGATACAAACATC 1235
QY 381 LeuHleIleProProGlnTrpLeuHleLeuValAlaAsnGlnTrpPheHleAspLysHle 400
Db 1236 CTGCACATCCCGCTCAGATTTGTGACCTTGTGGCTTAAGAAATCTTCAATGCAAGGAC 1295
QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
Db 1296 TCCCTGACGGAATTCGAGACAGCTCTTGTGACTTCTGTGAGATGTGTTTGTGTGTC 1355
QY 421 ProAlaLeuIleThrAlaArgTrpHleSarGAspAlaGlyAlaProValTrpPheTrpGlu 440
Db 1356 CCTGACATGATCAACAGCTGATATCAAGAGATGCTGTGCACTCTTCACTTATGAG 1415
QY 441 PheArgHleSarProGlnCySPheGlyAspTrpLysProAlaPheValLysAlaAspHle 460
Db 1416 TTTCGGCACCGGCTTATGAGTCTTTGAAGACAGAACCCGACTTTGTCAAGACCGACAC 1475
QY 461 AlaAspGlnValaArgPheValaPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480

D	b		1476	GCTGATGAAGTCCCGCTTTGTGTGGTGTCCTTCGAAAGGGGACATTGTATGTC	1535
O	y		481	GlucylalathrglugluglulysleuleuseraarglysmetMetyleYrTrpAlaThr	500
D	b		1536	GAAAGACCACGGAGGAGAGAATACTACTGAGCCGGAAGATGAAATACGCGGCTCAC	1595
O	y		501	PheIlaIrgrHrgrLyaaenProAnaGlyAsnAspleuSerLeuTPProAlaTYrAsnIeu	520
D	b		1596	TTTTGCTCGAACCGGGAATCCTTAATGGGAAGCACTGTCTGTGGGCCAGCTTAATATCTG	1655
O	y		521	ThrgluGlnTyrlaEugInIleuAspleuAsnMetSerIeugIyGIAnaArgLeuLysGluPro	540
D	b		1656	ACTAGACGTRACCTCCAGCTGSACTTGAAACATGAGCTCCGACAGAGACTCAAGAACCG	1715
O	y		541	ArgValAspPheTrprHrSerThrIleProleuIleuSeralAserAspMetLeuHis	560
D	b		1716	CGGGTGAATTTTGGACACAGCACCATCCCTGTATCTGTCTGCCTCGACATGCTCAC	1775
O	y		561	SerProLeuSerSerLeuThrPheluSerIeuleuGlnProPhePhePheCyAla	580
D	b		1776	AGTCCCTTTTCTTCTTAACCTTCTCTCTCTCTCCAGCCCTTCTTTTGTGCT	1835
O	y		581	Pro 581	
D	b		1836	CCT 1838	
RESULT 2					
A	DQ89093	ID	ADQ89093	standard; cDNA; 2158 BP.	
X	X	AC	ADQ89093;		
D	T	DT	21-OCT-2004	(first entry)	
X	X	DE	Human urological disorder related protein 53010 encoding cDNA SEQ:45.		
K	M	KM	urological disorder; uropathic; cytostatic; urinary incontinence;		
X	X	OS	benign prostatic hyperplasia; human; gene; ss.		
X	X	XX	Homo sapiens.		
F	H	Key	Location/Qualifiers		
F	T	CDS	96..1841		
F	T	FT	/*tag= A		
X	X	PN	/product= "urological disorder related protein 53010"		
X	X	PD	WO2004065576-A2.		
X	X	PD	05-AUG-2004.		
X	X	PF	14-JAN-2004; 2004WO-US000750.		
X	X	PR	15-JUN-2003; 2003US-0440318P.		
X	X	PR	04-FEB-2003; 2003US-0444783P.		
X	X	PR	27-MAR-2003; 2003US-0457901P.		
X	X	PR	08-MAY-2003; 2003US-0468775P.		
X	X	PR	19-MAY-2003; 2003US-0471614P.		
X	X	PR	16-JUN-2003; 2003US-0478742P.		
X	X	PR	18-JUL-2003; 2003US-0488522P.		
X	X	PR	30-JUL-2003; 2003US-0491156P.		
X	X	PR	02-SEP-2003; 2003US-0495949P.		
X	X	PR	26-SEP-2003; 2003US-0506332P.		
X	X	PA	(MILL-) MILLENNIUM PHARM INC.		
X	X	PI	Karicheti V, Silos-Santiago I, Eliasof SD;		
X	X	WI	MP1; 2004-562167/54.		
X	X	DR	P-PSDB; ADQ89094.		
X	X	PT	Use of polypeptides related to urological disorders, e.g. 44390, 54181,		
X	X	PT	211 or for identifying a compound capable of treating a urological		

PT	disorder or identifying and treating a subject having a urological disorder.
XX	
PS	Claim 1; SEQ ID NO 45; 542pp; English.
CC	The present invention describes the use of polypeptides related to
CC	urological disorders for identifying a compound capable of treating a
CC	urological disorder, identifying a subject having a urological disorder,
CC	or treating a subject having a urological disorder. Also described: (1) a
CC	method for identifying a compound capable of treating a urological
CC	disorder; (2) a method for identifying a subject having a urological
CC	disorder; and (3) a method for treating a subject having a urological
CC	disorder. The compound has uropathic and cytostatic activities. The
CC	polypeptides related to urological disorders are useful for identifying a
CC	compound capable of treating a urological disorder, identifying a subject
CC	having a urological disorder, or treating a subject having a urological
CC	disorder. Disorders include urinary incontinence and benign prostatic
CC	hyperplasia. The present sequence encodes a human urological disorder
CC	related protein, which is used in the exemplification of the present
CC	invention.
SQ	Sequence 2158 BP, 515 A, 559 C, 509 G, 575 T, 0 U, 0 Other;
Alignment Scores:	
Pred. No.:	4, 71e-282 Length: 2158
Score:	3079.00 Matches: 581
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatch: 0
Query Match:	100.00% Indels: 0
DB:	13 Gaps: 0
US-10-023-515-2 (1-581) x AD089093 (1-2158)	
DY	1 MetProGInGLyLeuThrSerSerAlaSerGlnTrpCySphePheLeuIleLeuGlnPro 20
Db	96 ATGCCACAGAGACTTAACTTCATCTGCCTCACAAATGATCTTTTCTCGATTCTCAAGCCC 155
DY	21 LeuLeuGLYHiaArgGLnTrpGLyIyrThrGLyProSerAlaGluGLyProGlnIlaArgAsn 40
Db	156 CTGTGGGACACAGACAAGTGGGAAAACTGGGCTTGTGTGAAGGGCCACAGAGAAC 215
DY	41 ThrArgLeuGLYTrpIleGlnGLyIyrArgGlnValThrValIleuGLySerProValProVal 60
Db	216 ACCAGGCTGGATGANTTACGGGACAGCAAGTACTGTGCTGGAGAAGCCCTGTGCTGTG 275
DY	61 AsnValPheLeuGLYValProPheAlaIalProPoleuGLySerLeuArgPheThrAsn 80
Db	276 AACGTTCCTCGAAGTCCCTTGTCTGCTCCCGCGTGAGATCCCTGCGATTACGAAAC 335
DY	81 ProGlnProAlaSerProTrpAspAsnLeuArgGLuAlaThrSerTYrProAsnLeuCys 100
Db	336 CCGCACCTGCATCCGCTCGGATTAACCTGGAGAAAGCACTCCTACCTCAATTTCG 395
DY	101 LeuGlnAsnSerGLuTrpLeuLeuLeuAspGlnHisMetLeuLysValHisTYrProLys 120
Db	396 CTCACAACTCAGAGGCGCTCTTAAATCAACATGCTCAAGGTGCATTACCAGAAA 455
DY	121 PheGLYValSerGLuAspCYLeuTYrLeuAsnIleTYrAlaProAlaHisAlaAspThr 140
Db	456 TTTCGAGCTGCAGAGACTGCTTCACTGAACATTAAGCGCTGGCCACGCCGATACA 515
DY	141 GlySerTyLeuLeuProValLeuValTrpPheProGLyGLyAlaPheLYrThrGLySerAla 160
Db	516 GGCTCAACATCCCGCTTGTGTGTGTCCACAGAGGTGCTTCAAGACTGCGCTCAGC 575
DY	161 SerTllePheAspGLYSerAlaLeuAlaIalTYrGluAspValIleuValIalValGln 180
Db	576 TCATCTTTGAATGGGCGCGCTGCTGCCTATAGGACGCTGCTGTGTGCTGCAG 635
DY	181 TYArgLeuGLYIlePheGLYPhePheThrTrpIleAspGlnHisAlaProGLYAsnTrp 200
Db	636 TACCGCTAGGAATATTGGTTTCTTCAACCATGGATCAGCATGGCTCGGGGAATGG 695

QY	201	AlaPheIysAepGlnValAlaIalaIeuSerTrpValGlnIysAsnIleGlnPhePheGly	220
Db	696	GCCTTCAGAGACCAAGGTGGCTGCTGCTGCTGGGTCAGAGAAACATGACTTCTTGCT	755
QY	221	GlyAsPProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSer	240
Db	756	GGGAGCCCAAGCTCTGTGACATCTTTGGCGAGTCCGGGAGGCCATAGGTTCTACGT	815
QY	241	IeuIleIeuSerProMetAlaIysGlyIeuPheIleIysAlaIleMetGluSerGlyVal	260
Db	816	CTTATATCTGTCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGGAGATGGGGTG	875
QY	261	AlaIleIleProTyrlleGlnIuAlaIleAspTrpGlyIysSerGluAspIeuGlnVal	280
Db	876	GCCATCATCTTACCTTACCTGGAGGCCCATATATATGAGAAATGAGACCTGGAGGTGGTT	935
QY	281	AlaHlePheCysGlyIAsnAsnAlaSerAspSerGluAlaIeuIeuAspGlySleuAsgTrp	300
Db	936	GCACATTTCTGTGTACAAATGCCGTACAGCTCTGAGGCCCTGCTGAGGTCTCGAGACA	995
QY	301	IysPProSerIysGlnIleuIeuThrIeuSerGlnIysThrIysSerPheThrArgVal	320
Db	996	AAACCTTCAGAGAGACTGTACCTCAGCCAGCAAAACAAAGCTTTCACCTCGAGGTGT	1055
QY	321	AspGlyAlaPhePheProAsnGluProIeuAspIeuSerGlnIysAlaPheIysAla	340
Db	1056	GATGGTGGCTTCTTCTCTTAATGAGCCTCTGATCTATGTCTCAGAAAGCATTTAAAGCA	1115
QY	341	IleProSerIleIleGlyValAsnAsnHleGluCysGlyPheIeuIeuProMetIysGlu	360
Db	1116	ATTCCTTCATCATTCGAGGTCAATAAACAGATGGCTTCTGCTGCTCATGAAAGAG	1175
QY	361	AlaProGluIleIeuSerGlySerAsnIysSerIeuAlaIeuHleIeuIleGlnAsnIle	380
Db	1176	GCTCTGTGAGATCCTCAAGTGGCTCAACAACATCCCTTGCCCTCATCTGATACAAACATC	1235
QY	381	IeuHleIleProProGlnItyrlleuHleIeuValAlaAsnGluTyrlPheHleAspIysHle	400
Db	1236	CTGACATCCCGCTCGATGTATTGACCTTGTGGCTAAAGAAATACCTTCATGACAGAC	1295
QY	401	SerIeuThrGlnIleArgAspSerIeuIeuAspIeuIeuGlyIAspValPhePheValVal	420
Db	1296	TCCCTGACTAAATCCGAGACATCTTCTTGACTTCTTGAGATGTGTCTTGTGGTCTC	1355
QY	421	ProAlaIeuIleThrAlaArgTyrlHleAspAlaGlyAlaProValTyrlPheTyrlGlu	440
Db	1356	CTCGACTGATCAACACTCGATTCACAGAGATGGTGGACCTGTCTATCTTATGAG	1415
QY	441	PheArgHleAspProGlnCysPheGluAspTrpIlyAspAlaPheValIysAlaAspHis	460
Db	1416	TTTCGGGACCGGCTCATGTCTTTGAGACAGAGACCGGCTTTGTCAAGCCGACAC	1475
QY	461	AlaAspGluValArgPheValPheGlyGlyAlaPheIeuIysGlyAspIleValMetPhe	480
Db	1476	GCATGAGAAATCGCTTGTGTTCGGTGGTGCCTTCGAAAGGGGACATTTGTTATCTTC	1535
QY	481	GlnGlyAlaIleThrGlnGlnGlnIuIysIeuIeuSerArgIysMetMetIyrylTrpAlaThr	500
Db	1536	GAAAGGCCACCGAGAGAGAGAAATTACTGAGCCGAGAAAGATGAATATCTGGGTACCC	1595
QY	501	PheAlaArgThrGlyAsnProAsnGlyIAsnAspIeuSerIeuTrpProAlaTyrlAsnIeu	520
Db	1596	TTTGTCTCGAAACCGGGAATCTTAATGGAGACGCTGTCTGTGGCCACTTAAATCTTG	1655
QY	521	ThrGlnGlnTyrlleuGlnIeuAspIeuAsnMetSerIeuGlyIuIArgIeuIysGluPro	540
Db	1656	ACTGAGCAGTACTCTCACTGAGACTTGAAATGAGGCTCGAGACAGACATCAAAAGAACCG	1715
QY	541	ArgValaAspPheTrpThrSerThrIleProIeuIleIeuSerAlaSerAspMetIeuHis	560
Db	1716	CGGGTGGATTTTGGACAGACCATCCCTCATCTGTCTGCTCTCCAGATGCTCCAC	1775
QY	561	SerProIeuSerSerIeuThrPheIeuSerIeuIeuGlnProPhePhePheCysAla	580

Db	1776	AGTCTCTTTTCCTTAACCTTCCTCCTCCCTCACGCTTCTTTCTTTGGCT
Oy	581 Pro 581	
Db	1836 CCT 1836	
RESULT 3		
ID	AB086169 standard; DNA; 1728 BP.	
XX	AB086169;	
DT	10-SEP-2002 (first entry)	
DE	Novel human gene. SEQ ID 40.	
XX		
KW	Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;	
KW	nootropic; neuroprotective; immunosuppressive; haemostatic;	
KW	antiinflammatory; cardiant; anticulcer; vincide; antithyroid;	
KW	cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;	
KW	wound healing disorders; atherosclerosis; Parkinson's disease;	
KW	Alzheimer's disease; autoimmune disorder; haematopoietic disorder;	
KW	inflammation; neoplastic disease; nervous system disorder;	
KW	cardiovascular disorders; pancreatitis; respiratory disorder;	
KW	hyperproliferation; systemic autoimmune disease; hyper-immunity;	
KW	developmental abnormality; gastrointestinal ulceration; neuropathy;	
KW	haematological abnormality; gastrointestinal ulceration; hypothyroidism;	
KW	thyroid disorder; hypothyroidism; brain damage; colitis;	
KW	come photo-transduction deficiency; neurological disease; stroke;	
KW	angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;	
KW	trachea; thymus; lymph node; muscular system; obesity; anorexia;	
OS	growth abnormality; precocious puberty; gene; ss.	
XX		
XX	Homo sapiens.	
XX	MO200250105-A1.	
XX		
PD	27-JUN-2002.	
XX		
PF	17-DEC-2001; 2001WO-US049232.	
XX		
PR	19-DEC-2000; 2000US-0256710P.	
PR	20-DEC-2000; 2000US-0257048P.	
PR	09-JAN-2001; 2001US-0260482P.	
PR	30-JAN-2001; 2001US-0264922P.	
PR	06-FEB-2001; 2001US-0266797P.	
PR	19-MAR-2001; 2001US-0276988P.	
PR	04-APR-2001; 2001US-0281355P.	
PR	08-MAY-2001; 2001US-0289622P.	
XX		
PA	(SMK) SMITHLINE BEECHAM CORP.	
PA	(SMK) SMITHLINE BEECHAM PLC.	
PA	(GLAX) GLAXO GROUP LTD.	
XX		
PI	Agarwal P, Birkeland M, Cogswell JP, Kabnick KP, Lai Y,	
PI	Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q,	
DR	WPI; 2002-508784/54.	
XX	P-PEDB; ABP61004.	
PT	Secreted proteins and polynucleotides useful as vaccines for preventing	
PT	or treating various diseases e.g. cancer, wounds, atherosclerosis,	
PT	Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.	
PS	Claim 2(a); Page 251; 335pp; English.	
XX		
CC	The invention relates to an isolated polypeptide with signal sequences	
CC	which allow it to be secreted extracellularly or membrane associated. The	
CC	activity of polypeptides of the invention may be described as,	
CC	cytostatic, vulnary, antiarteriosclerotic, antiparkinsonian, nootropic,	
CC	neuroprotective, immunosuppressive, haemostatic, antiinflammatory, anti-	
CC	cardant, anticulcer, vincide, antithyroid, cerebroprotective, etc.	

CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC wound in the treatment, or as a vaccine in the prevention of, cancer,
 CC inflammation, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, hematopoietic disorder,
 CC cardiovascular disorders, nervous system related disorders and
 CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
 CC developmental abnormality, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, angiodenesis,
 CC ovalation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
 CC of the invention

XX Sequence 1728 BP; 367 A; 483 C; 437 G; 441 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,3e-266	Length:	1728
Score:	2915.00	Matches:	553
Percent Similarity:	97.54%	Conservative:	3
Best Local Similarity:	97.02%	Mismatches:	6
Query Match:	94.67%	Indels:	8
DB:	6	Gaps:	1

US-10-023-515-2 (1-581) x ABQ86169 (1-1728)

```

QY 12 TrpCyvphelheleuileuGlnProleuileuGlnleuGlntrpGlyleuThrgly 31
DB 40 TGGGCTACTGGGCTCTTGCACACCC-----ACCAAGG 75
QY 32 ProSerAlaGlyGlyProGlnArgAnthArgleuGlyTrpIleGlnGlyleuVal 51
DB 76 CCTTGTCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCAATC 135
QY 52 ThrValleuGlySerProValProValAsnValPheleuGlyValProheAlaIlePro 71
DB 136 ACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTCTCTGAGAGTCCCTTGTCTGTCTCC 195
QY 72 ProleuGlySerleuArgPheThrAsnProGlnProAlaSerProTrpAsnleuArg 91
DB 196 CCGCTGGGATCCCTCGATTTCAGAACCCGACCGCTGCATCCCTGGGATACCTTGGCA 255
QY 92 GluAlaThrSerTrpProAsnleuCyseuGlnAsnSerGlyTrpIleuLeuAspGln 111
DB 256 GAAGCCACCTCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTCTTAGATCA 315
QY 112 HisMetleuValHisTrpProValPheGlyValSerGluAspCyseuLeuAsn 131
DB 316 CATATGCTCAAGGTGATTAATCCGAAATTCGAGGTCTCAGAACTGCTCTTACCTGAAC 375
QY 132 IleTrpAlaProAlaHisAlaAspThrglySerIleuProValleuValTrpPhePro 151
DB 376 ATCTATGGCCCTGCCACCGCGATACAGGCTCCAACTCCCGTCTTGGTGTGGTTCCCA 435
QY 152 GlyIleAlaPheLeuThrglySerAlaSerIlePheAspGlySerAlaLeuAlaIleTrp 171
DB 436 GGAAGTGGCTTCAAGACTGGCTCAGGCTCATTTGATGGGTCCGCTGCTGCTGCTAT 495
QY 172 GluAspValleuValValValGlnTrpArgleuGlyIlePheGlyPhePheThrTrp 191
DB 496 GAGGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
QY 192 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLeuAspGlnValAlaIleAspSerTrp 211
DB 556 TGGGATCAGCAAGCTCTCGGGAACTGGGCTTCAAGACCAAGTGGCTCTGTCTGTGG 615
QY 212 ValGlnValAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu 231
DB 616 GTCCAGAGAAACATGAGTTCTTGGTGGGAGACCCAGGCTGTGATCACTTTTGGCGAG 675

```

```

QY 232 SerAlaGlyAlaIleSerValSerSerleuIleLeuSerProMetAlaGlyGlyleuPhe 251
DB 676 TCCGGGGAGGACCAATAGTCTTCTAGTCTTAAGTGTCCCATGGCCAAAGGCTTATTC 735
QY 252 HisValAlaIleMetGluSerGlyValAlaIleIleProTyrlleuGluAlaHisAspTrp 271
DB 736 CCAAAAGCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGGCCCATGATTTAT 795
QY 272 GluValSerGluAspLeuGlnValValAlaHisPheCyseuGlyAsnAsnAlaSerAspSer 291
DB 796 GAAAGAGTGAAGACCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 855
QY 292 GluAlaLeuLeuArgCyseuArgThrglyAspProSerIleGlyleuLeuThrglySerGln 311
DB 856 GAGGCTCTGTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 915
QY 312 TyrlleuValSerPheThrArgValAlaAspGlyAlaPhePheProAsnGluProleuAsp 331
DB 916 AAAACAAAGCTTTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 975
QY 332 LeuLeuSerGlnValAlaPheValAlaIleProSerIleIleGlyValAsnAsnHisGlu 351
DB 976 CTATGTGTCAAGAAAGCATTTAAACCAATTCCTTCATCATGAGTCAATACCAAGAG 1035
QY 352 CyseuGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLysSer 371
DB 1036 TGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1095
QY 372 LeuAlaLeuHisValleuIleGlnAsnIleleuHisIleProProGlnTyrlleuHisVal 391
DB 1096 CTGGCCCTCATCTGATCAAAACATCTGCAGCATCCGCTCAGATATTGACACTGTGTG 1155
QY 392 AlaAsnGlyTrpPheHisAspLysHisSerleuThrglyIleArgAspSerleuLeuAsp 411
DB 1156 GCTAATGAATATCTTCATGACAAAGCACTCCCTGACTGAATTCGAGACAGTCTTGTGAC 1215
QY 412 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrlleAspAsp 431
DB 1216 TTGCTTGAAGATGTCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1275
QY 432 AlaGlyAlaProValTyrllePheTyrllePheArgHisArgProGlnCyseuPheGluAspTrp 451
DB 1276 GCTGTGACACCTGTCTACTTATGAGTTTGGGACACGGCTCAGTGTCTTGAAGACAG 1335
QY 452 LysProAlaPheValLysAlaAspHisAlaAspGlyValArgPheValPheGlyGlyAla 471
DB 1336 AAGCAGCTTTTGTAAAGCCACCACTGTGAAGTCCCGCTTGTGTGTGTGTGTGTGTGT 1395
QY 472 PheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGlyValLeuLeuSer 491
DB 1396 TTCTGAAAGGGGACATTTGTTATGTTCCAAAGGCCACGAGAGAGAGATTAAGTAC 1455
QY 492 ArgLysMetMetLysTyrlleTrpAlaThrPheAlaArgThrglyAsnProAsnGlyAsnAsp 511
DB 1456 CGGAAGATGATGAATATCGGGCTACCTTGTGTGACACGGGAATCTTAATGGAGAACAG 1515
QY 512 LeuSerleuTrpProAlaTyrlleAsnleuThrglyGlnTyrlleuGlnleuAspLeuMet 531
DB 1516 CTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1575
QY 532 SerleuGlyGlnArgleuLeuGluProArgValAspPheTrpThrSerThrIleProleu 551
DB 1576 AGCTTGGACAGAGCTCAAGAACCGGGGTGGATTTTGGACAGACCATTCCTCTG 1635
QY 552 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerleuThrPheLeuSerleu 571
DB 1636 ATCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1695
QY 572 LeuGlnProPhePhePhePheCysAlaPro 581
DB 1696 CTCCAGCTTTCTTTTCTTTTGTGTCTCT 1725

```

RESULT 4

ID	AAD40574	bctandard; cDNA; 2232 BP.
XX	AC	AAD40574;
XX	DT	30-OCT-2002 (first entry)
DE	Human drug metabolizing enzyme (DME-10) cDNA.	
KM	Human; drug metabolizing enzyme; autoimmune; inflammatory disorder.	
KW	acquired immunodeficiency syndrome; AIDS; atherosclerosis; psoriasis;	
KW	proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;	
KW	asthma; neurological disorder; Alzheimer's disease; Huntington's disease;	
KW	dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;	
KW	renal screening; endocrine disorder; conjunctivitis; glaucoma; cataract;	
KW	renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;	
KW	anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;	
KW	giltire; gastrointestinal disorder; gene therapy; virulence; anticoagulant;	
KM	anticonvulsant; nootropic; enzyme; DME-10; gene; 88.	
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..1929
FT		/*tag= a
FT		/product= "Human DME-10"
FT	sig_peptide	1..150
FT		/*tag= b
FT	mat_peptide	151..1926
FT		/*tag= c
FT		/product= "Mature human DME-10"
PN	WO200246426-A2.	
XX		
PD	13-JUN-2002.	
XX		
PX	04-DEC-2001; 2001MO-US047429.	
PR	08-DEC-2000; 2000US-0254308P.	
PR	15-DEC-2000; 2000US-0256189P.	
PR	21-DEC-2000; 2000US-0257113P.	
PR	19-JAN-2001; 2001US-0262706P.	
PR	02-FEB-2001; 2001US-0266020P.	
PA	(INCY-) INCYTE GENOMICS INC.	
PI	Sanjanwala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ,	
PI	Lee EA, Ding L, Hafalia AJA, Tang YT, Yue H, Tribouley CM, Lu DM,	
PI	Lal PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AR, Lu Y,	
PI	Ison CH,	
DR	WPI; 2002-519668/55.	
PT	P-PSDB; AAE25025.	
PT	Novel human drug metabolizing polypeptide, useful in diagnosis,	
PT	prevention or treatment of autoimmune/inflammatory, cell proliferative,	
PT	neurological, developmental, endocrine, metabolic and gastrointestinal	
XX	disorders.	
PS	Claim 78; Page 167-168; 169pp; English.	
CC	The invention relates to an isolated human drug metabolizing enzyme (DME-	
CC	and its nucleotide. DME is useful for diagnosing, treating or preventing	
CC	disorders associated with aberrant expression of DME, where the disorders	
CC	are selected from autoimmune/inflammatory disorder such as acquired	
CC	immunodeficiency syndrome (AIDS), asthma, atherosclerosis, psoriasis,	
CC	incretin deficiency syndrome (AIDS), dementia, Huntington's disease;	
CC	cirrhosis; a cell proliferative disorder such as arteriosclerosis,	
CC	cirrhosis, hepatitis, and cancer; a neurological disorder such as	
CC	Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;	
CC	a developmental disorder such as renal tubular acidosis, epilepsy,	
CC	anaemia; an endocrine disorder such as adenoma, thrombosis and infections	
CC	anaemia; an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic	
CC	disorder such as cystic fibrosis, diabetes and goitre; a gastrointestinal	


```

OY 291 SerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer 310
DB 1054 TCTGAGGCGCTCTGAGGCTGCTGAGAGCAAAACCTCCAGAGAGCTGCTGACCTCAGC 1113
OY 311 GlnLysThrLysSerPheThrArgValValAaspGlyValAphePheProAsnGluProLeu 330
DB 1114 CAGAAACCAAAAGCTCTTCACTCGAGGTGATGGGCTTTCTTCTTCTTAATAGGCTCTTA 1173
OY 331 AspLeuLeuSerGlnLysAlaAphelysAlaIleProSerIleIleGlyValAsnAsnHis 350
DB 1174 GATCTATTGCTCAGAAACCAATTAAAGCAATCTTCCATCATCGAGCTCATATACCAAC 1233
OY 351 GluGlyGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLys 370
DB 1234 GAGTGTGGCTTCTGCTGCTATGAAAGAGGCTCTGAGGGTCTCTAGTGGCTCCACAAAG 1293
OY 371 SerLeuAlaLeuHisAlaLeuIleGlnAsnIleLeuHisIleProProGlnThrLeuHisLeu 390
DB 1294 TCCCTTGGCCCTCATCTGATACAAACATCCCGACATCCCGCTCAGTATTGGACCTT 1353
OY 391 ValAlaAsnGluThrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeu 410
DB 1354 GTGGCTAAATGAATACTTCCATGACAAAGCACTCCCTGACTGAATCCGAGACAGCTTCTG 1413
OY 411 AspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgThrHisArg 430
DB 1414 GACTTGCTTGGAGAGTGTTCTTGTGTGCTGCTGACTGATCAAGCTGATATACAA 1473
OY 431 AspAlaGlyAlaProValThrPheThrGluPheArgHisArgProGlnCysPheGluAsp 450
DB 1474 GATGCTGGGACACCTGTACTTCTATGAGTTTCGGACCGGCTCAGTGGCTTTGAAAGC 1533
OY 451 ThrLysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGly 470
DB 1534 AGGAAAGCGGCTTTGTCAAGCCACCAAGCTGATGAAGTCCGCTTGTGTGCGTGGT 1593
OY 471 AlaPheLeuLysGlyAspIleValMetPheGluGluValAlaThrGluGluGluLysLeuLeu 490
DB 1594 GCTTCTCGAAGGGGACATGTGTATGTGCAAGGACCAAGAGAGAGAGAAAGTACTG 1653
OY 491 SerArgLysMetMetLysIleThrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 510
DB 1654 AGCCGGAAGATGATGAATACTGGGCTACCTTGTCCGACCGGGAATCTTAATGGAAAC 1713
OY 511 AspLeuSerLeuThrProAlaThrAsnLeuThrGluGlnThrLeuGluLeuAspLeuAsn 530
DB 1714 GACCTGTTTCTGTGGCCAGCTTATATCTGACTGAGAGTACTCCAGCTGAGACTTGAAC 1773
OY 531 MetSerLeuGlyGlnArgLeuLysGluProArgValAaspPheThrThrSerThrIlePro 550
DB 1774 ATGAGCCCTCGACAAAGACTCAAAAGAACCGGCGGTGGATTTTGGACCGACCAACATCCCC 1833
OY 551 LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer 570
DB 1834 CTGATCTCTGTGCTCTCGACATGCTCCACAGTCTCTTCTTCTTAACCTTCTCTCT 1893
OY 571 LeuLeuGlnProPhePhePhePheCysAlaPro 581
DB 1894 CTCTCCAGACCTTCTTTTCTTTTGTGTCTCT 1926

```

RESULT 5

ACN42766
ID ACN42766 standard; cDNA; 2229 BP.

ACN42766;
AC ACN42766;
DT 18-NOV-2004 (first entry)

XX Human diagnostic and therapeutic polynucleotide SEQ ID NO:1641.
XX Human diagnostic and therapeutic polynucleotide;
XX 89; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.

```

XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US028227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCYTE) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
XX PI Hartschorne TM, Suchorolski MT, Altus CM, Plets SJ, Elder LV,
XX PI Mooney EM, Delegeane AM, Panesar IS, Banyille SC, Reddy TP,
XX PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Geregin EH,
XX PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL,
XX PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitti UA, Kilton BS,
XX PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D,
XX PI Patury S, Shi X, Suarez CJ,
XX DR WPI; 2004-329368/30.
XX DR P-PSDB; ABW84114.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX PT in diagnosing a condition, disease or disorder associated with human
XX PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX PS in gene mapping.
XX PS Claim 1, Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
XX CC selected from one of the 2722 sequences defined in the specification. A
XX CC polynucleotide of the invention may have a use in gene therapy. The human
XX CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX CC used to diagnose a particular condition, disease or disorder associated
XX CC with human molecules, e.g. cell proliferative disorders,
XX CC autoimmune/inflammatory disorder, developmental disorder, endocrine
XX CC disorder, neurological disorders, gastrointestinal disorders, or
XX CC infections caused by virus, bacteria, fungi or parasite. The dithp
XX CC molecules may also be used in genetic mapping, in identifying individuals
XX CC from minute biological samples, in detecting single nucleotide
XX CC polymorphisms, as molecular weight markers, and for somatic or germ-line
XX CC gene therapy. The present sequence represents a dithp polynucleotide of
XX CC the invention. Note: The sequence data for this patent is not represented
XX CC in the printed specification, but was obtained in electronic format
XX CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 2229 BP; 503 A; 585 C; 547 G; 594 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.: 6,176-263 Length: 2229
Score: 2878.00 Matches: 549
Percent Similarity: 99.82% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 93.47% Indels: 1
DB: 13 Gaps: 0

US-10-023-515-2 (1-581) x ACN42766 (1-2229)

```

OY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyThrIleGlnGlyGln 50
DB 274 GGGCCCTTCTGCTGAAAGGCGCAAGAGAAACACAGGCTGGATGATTCAGAGGCAAGCAA 333
OY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIle 70
DB 334 GTCACTGTCGTGGAAAGCCCTGTGCTGTAACGTTCTCTCGAGTCCCTTCTCTGCT 393
OY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAspAsnLeu 90
DB 394 CCCCCTGGGATCCCTCGATTTCAGAACCGGACGCTGATCGCCCTGGGATTAACCTTG 453

```


QY	91	ArgGluValaIthSerTyProAsnLeuCyLeuGlnAsnSerGluTrpLeuLeuAsp	110
Db	454	CGAAGAACCACTCCCTAACCCTAATTGGCCCTCGAAGCTCAGAGTGGCTCTTGAT	513
QY	111	GlnHsMeLeuLeuValHsIeTyProIyAsnHeGlyValSerGluAspCyLeuTyLeu	130
Db	514	CAACATATGCTCAAGTGCATTACCCGAATTCGAGTGTCAAGAACCTGCTCTACTG	573
QY	131	AsnIeTyraIProAlaHsAlaAspThrGlySerIySerIyLeuProValLeuValTrpPhe	150
Db	574	AACATCTATGGCGCCGCCAAGCGATACAGGCTCCAAAGCTCCCCGCTTGGTGGTTC	633
QY	151	ProGlyGlyAlaPheIySerThyGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa	170
Db	634	CCAGAGGTGGCTTCACAGACTGGCTCAGCCCTCCATCTTATATGGGTCCGCCCTGGCTGCC	693
QY	171	TyrGluAspValLeuValValValGlnTyraArgLeuGlyIlePheGlyPhePheThr	190
Db	694	TATAGAGACGTCGTGGTTGTGTGTGTCCAGTACCGGCTAGGAATATTGGTTCTC-AGC	752
QY	191	ThrTrpAspGlnHsAlaProGlyAsnTrpAlaPheIyAspGlnValAlaAlaLeuSer	210
Db	753	ACATGGGATCAGCAATGCTCCGGGAACTGGGCTTCMAAGACAGAGTGGCTGTGCTCC	812
QY	211	TrpValGlnIyAsnIleGluPhePheGlyGlyAspProSerSerAlaThrIlePheGly	230
Db	813	TGGGTCCAGAAACATCCAGATCTTGCTGGTGGGAAACCCAGCTGTGTGACATCTTTGGC	872
QY	231	GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIyGlyLeu	250
Db	873	GAGTCGGCGGAAACCATATAGTGTTCAGTCTTATAGTCTGTCCATGGCCAAAGGCTTA	932
QY	251	PheHsIyValaIleMetGluSerGlyValAlaIleIleProTyrlLeuGluAlaHsAsp	270
Db	933	TTCCCAAGACCATCATGAGAGTGGGGTGGCCATATCCCTTACCTGGAGGCCCATAT	992
QY	271	TyrGluIySerGluAspLeuGlnValValAlaHsPheCyGlyAsnAsnAlaSerAsp	290
Db	993	TATGAAAGAGTGAGGACCTCGACAGTGGTTGCACATTTCTGTGTACATATCCGTCAAC	1053
QY	291	SerGlnAlaLeuLeuArgCyLeuAspGlyTrpProSerIyGlnLeuLeuThrLeuSer	310
Db	1053	TTCGAGGCCCTGTGAGGTGCTGAGGACAAACCCCTCCAGAGACTCTACCTCAGC	1112
QY	311	GlnIySerThySerPheThrArgValAlaAspGlyAlaPhePheProAsnGluProLeu	330
Db	1113	CGAAGAACCAAGTCTTCTCAGAGTGTGATGGTCTTCTTCCATAGAGCTCTA	1172
QY	331	AspLeuLeuSerGlnIyValaPheIyValaIleProSerIleIleGlyValaAsnHsIy	350
Db	1173	GATCTATTGTCTCAAGAACATTAAAGCAATCTTCCATCATCGGAGTCAATACAC	1232
QY	351	GluCyGlyPheLeuLeuProMetIyGlnAlaProGluIleLeuSerGlySerAsnIyS	370
Db	1233	GAGTGTGGCTCTCTGCTGCTATGAAGAGGTCTTGAGATCTCAGTGGTCTCAACAG	1292
QY	371	SerLeuAlaLeuHsIleLeuIleGlnAsnIleLeuHsIleProProGlnTyrlLeuHsIe	390
Db	1293	TCCCTTGGCCCTCACTGATACAAACATCTCGACATCCCGGCTCAGTATTGCACTT	1352
QY	391	ValAlaAsnGluTyrlPheHsIaAspIyHsIaSerLeuThyGluIleAspAspSerLeuLeu	410
Db	1353	GTGGCTATGAATCTTCATGACAGACCTCCCTGACTGAATACCGAAGACGCTTCTG	1412
QY	411	AspLeuLeuGlyAspValPhePheValaIaProAlaLeuIleThrAlaArgTyrlHsIaArg	430
Db	1413	GACTTGCTTGGAGATGTGTCTTGTGTGCTCCTGCACTGATCAGCTGCATTCACAG	1472
QY	431	AspAlaGlyAlaProValTyrlPheTyrlGluPheArgHsIaArgProGlnCyPheGluAsp	450
Db	1473	GATCTGGTGCACCTGTCTACTTATAGAGTTTGGGACCGGCTCAGTGTCTTGAAGAC	1532

Oy		451	ThrlrYsrPrOAlaPheValIylValaaSPHlSaLaasPgUValArgPheValPheGlyGly	470
Db		1533	ACGAAGCGACCTTTGTTCAAAGCCGCACACGGCTGATGAAGTCCGCTTGTTGGTGGT	1592
Oy		471	AlaPheLeuIylAspIleValMetPheGluGlyValAthrGluGluGluLysLeuLeu	490
Db		1593	GCCCTCCTTAAGGGGCACTGTTATTGTTCCAAAGGCCACGAGAGAGAAGTACTG	1652
Oy		491	SerArgLysMetMetIylsYrTrpAlaThrPheaIaaArgThrGlyAsnProAsnGlyAsn	510
Db		1653	AGCCGGAAGATGATGAATAACTGGGCTACTTGTGCTCCAGCCGGAAATCCTAAAGGAAAC	1712
Oy		511	AspLeuSerLeuTrpProAlaYrTrnLeuThrGluGlnTyrlLeuGlnLeuAspLeuAsn	530
Db		1713	GACCTGCTCTGTGGCCACCTTAATCATGACTGAGCGAGTACCCTCCAGCTTGAACTTGAAAC	1772
Oy		531	MetSerLeuGlyGlnArgLeuIylAspArgValAspPheTrpMserThrIlePro	550
Db		1773	ATGAGCTCGCGACAGAGACTCAAAGAACCGCGGGTGGATTTTGGACGACGACCATCCC	1832
Oy		551	LeuIleLeuSerIlaSerAspMetLeuHisSerProLeuSerSeriLeuThrPheLeuSer	570
Db		1833	CTGATCTCTGTGCTCCGACATGCTCCACAGTCTCTCTTCTTCTTAATTCTCTCT	1892
Oy		571	LeuLeuGlnProPhePhePheCysAlaPro	581
Db		1893	CTCCTCCAGCCTTTCTTTTTCTTTGTGCTCCT	1925
RESULT 6				
ID	ADB62095		standard; cDNA; 2092 BP.	
XX				
XX	ADB62095;			
XX				
DT	04-DEC-2003	(first entry)		
De			Human cDNA encoding clone BRAWH20021910.	
XX				
KW	Human; ss; gene; pharmaceutical; diagnostic; gene therapy;			
KM	tissue regeneration; cell regeneration; membrane protein;			
KW	signal transduction-related protein; transcription-related protein;			
KX	osteoporosis; neurological disease; cancer; tumour.			
OS	Homo sapiens.			
XX				
FH	Key	Location/Oualifiers		
FT	CDS	394..1803		
FT		/tag= a		
FT		/product= "Clone BRAWH20021910 protein"		
PN	EP1308459-A2.			
XX				
PD	07-MAY-2003.			
XX				
PF	28-MAR-2002; 2002EP-00007401.			
PR	05-NOV-2001; 2001JP-00379298.			
PR	25-JAN-2002; 2002US-00350978.			
XX				
PA	(HELI-) HELIX RES INST.			
XX	(REAS-) RES ASSOC BIOTECHNOLOGY.			
P1	Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehii S,			
P1	Yanamoto J, Isono Y, Hiro Y, Otsuka K, Nagai K, Irie R, Tamechika I,			
P1	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maeno Y,			
DR	WPI; 2003-450961/43.			
DR	P-PSD; ADB64065.			
XX				
XX	New polynucleotides and polypeptides, useful for developing a diagnostic			
PT	marker or medicines for regulation of their expression and activity, or			
PT	as targets of gene therapy.			
XX				

PS Claim 1, Page: 222pp, English.

XX The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.

XX Sequence 2092 BP: 464 A; 552 C; 497 G; 579 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7e-263	Length:	2092
Score:	2877.00	Matches:	551
Percent Similarity:	98.40%	Conservative:	2
Best Local Similarity:	98.04%	Mismatches:	2
Query Match:	93.44%	Indels:	7
DB:	10	Gaps:	2

US-10-023-515-2 (1-581) x ADB62095 (1-2092)

QY	24	HISARGGINTTPRGILYVETHRGLYPROSERALAGLUGLYPROGLINARGASNTHRARGLEU	43
DB	126	CACCTTAACTGG-----GGGCTTCTGCTGAAGGCCACAGAGAACCGAGGCTG	176
QY	44	GLYTPRIIEGLNGLYVSGINVALTHRVALLLEUGLYSERPROVALPROVALASNAVLPHE	63
DB	177	GGATGATTCACGAGCAAGCACTGCTGCTGGAGACCCCTGCTGTGAACCTGTTTC	236
QY	64	LEUGLYVALPROPHALALABROPROLEUGLYSERLEUARGPHETHRANBPROGLINPRO	83
DB	237	CTCGGAGTCCCTTGTGCTGCTCCCGCTGGGATCCCTGCGATTACGAAACCGGACGCT	296
QY	84	ALASERPROTPRAPHANLEUARGILVALATHRSERYTPROANLEU-----CY	100
DB	297	GCATGCGCCCTGGGATTAATCTTGCAAGAGCCACTCTTACCTTAATTGTAAAGACGAGTG	356
QY	100	ALEUGLINANSERGIUTRPLEULEULEUASPGINHSMETLEULYVVALHISLYRPROLY	120
DB	357	CCTCAAGAACTCAAGTGGCTGCTCTTAGATCAACAATGCTCAAGGTGCTATCCGAA	416
QY	120	SPHEGLYVALSERGIUAPCYBLEUTYRIEUBANILETYRALAPROVALAHISALASPTH	140
DB	417	ATTGGAGAGTCAAGAAAGCTGCTCACTGAACATCTATGCGCTGCGCCACGCGGATAC	476
QY	140	RGLYSERLYBLEUPROVALLEUVALTRPHEPROGLYGLYALAPHELYRTHGLYSERAL	160
DB	477	AGGCTCCAGAGCTCCCGCTTGCTGTGTGTCCACAGAGGTGCTTCAAGACGCGCTACGC	536
QY	160	ASERILEPHEAPGYSERVALALEUVALALATYRGILUAPVALLEUVALVALVALG	180
DB	537	CTCCATCTTTATAGGTGCTGCGCTGCTGCTATAGAGACGTGCTGTGTGTGTGTCTCA	596
QY	180	NTRYARGLEUGLYILEPHEGLYPHEPHETHRTRTPRAPHGINHISALAPROGLYASNTYR	200
DB	597	GTACCGGCTAAGAAATATTGTTTCTTCAACACATGGATGATGATGCTCGGGGAACTG	656

QY	200	PALAPHELYSAPSGINVALALALALEUSERTRYVALGINVYSAENILEGLUPHEPHEGL	220
DB	657	GGCTTCAAGAGCAAGAGTGGCTGCTCTCTGCTCCAGAAAGCACTGACTTCTTGG	716
QY	220	YGLYASPPROSERSERVALTHRIIEPHEGLYGLUSERALAGLYALALIESERVALSERSE	240
DB	717	TGGGAGCCCAAGCTGTGACCATCTTTGGCGAGTCCCGGGAGCCATAAGTGTCTTAG	776
QY	240	RLLEULEUSERPROMETALALYSGLYLEUPHEHISLYVALIEMETGLUSEGLYVA	260
DB	777	TCTTATCTGCTCTCCCAAGGCCMAAGGCTTATTCACAAAGCCATCATGAGAGTGGGT	836
QY	260	IALATILEPPOTRYLEUGLVALAHISAPTYRGILYSSERGIUAPLEUGLVALIVA	280
DB	837	GGCCATATCCCTTACCTGAGAGCCCAATATTAAGAAAGTAGAGACCTCCAGTGGT	896
QY	280	IALAHISAPHECYSGLYAENASNAIASERASPSERGIUVALLEULEUARGYLEUARGTH	300
DB	897	TGCACATTTCTGTGTAAACAATGCGTCAGACTTGAGGCCCTGCTGAGTGGCTGAGGAC	956
QY	300	RLYSPROSERLYSGILEULEULEUTHLEUSERGLINLYRTHLYSERTPHETHRARGVALVA	320
DB	957	AAACCCCTCCAGAGAGCTGACCCCTCAGCCAGAAACCAAGTCTTCACTGAGTGGT	1016
QY	320	IASPGLYVALPHEPHEPROANGLUPROLEUASPLEUSERGLINLYSALAPHELYVAL	340
DB	1017	TGATGTGCTTTCTTCTTAAGCTCTTAGATCTTATGTCTCAGAAAGCATTTAAAC	1076
QY	340	ALIEPSEERILEIEGLYVALASNAHISGLINUCYSGLYPHELEULEUPROMETLYSG	360
DB	1077	AATTCCTTCCATCANTCCGAGTCAATATGACAGAGTGGCTTCCGTGCTGCTATAGAGA	1136
QY	360	UVALPROGLIULEUSERGLYSERASNYLSERLEUALALEUHSILEUILEGINASNI	380
DB	1137	GGCTCTGAGATCTCCAGTGGCTCCACACAGTCCCTTCCCTCCATCGATACAAACAT	1196
QY	380	ELEUHSILEPPOPROGLINTRYRLEUHSILEUVALALAHANGIUTYRPHHISAPPLYSHI	400
DB	1197	CTTGACATCCCGCTCACTGATTTTGACATTTGTGGCTTAGTAAATCTTCCATGACAGCA	1256
QY	400	ASERLEUTHRGILUEARGASERLEULEUASPLEULEUGLYASPVALLPHEPHEVALVA	420
DB	1257	CTCCCTGACTGAATCCGAGACAGTCTTCTGGAATGAGTGTCTTGTGTGT	1316
QY	420	IPROALALEUILLERHALAAGTYRHISARGASPALAGLYALAPROVALTYRPHETRYGL	440
DB	1317	CCCTGCACTGATCAACAGCTCGATATACAGAGATGCTGTGCACTGTCTACTTATAGA	1376
QY	440	UPHEHIGHTHARGPPROGLINCYSPHEGLUASPTHRLYSPROVALPHEVALLYVALASPHI	460
DB	1377	GTTCGGCAACGGCTCAAGTGTGTGTAAGACACGAAAGCGGCTTTGTCAAGCCGACCA	1436
QY	460	SALASAPGILUVALARGPHEVALPHEGLYGLYALAPHELEULYSGLYASPVILLEVALMECPH	480
DB	1437	CGCTGATGAAGTCCGCTTGTGTGTGCTGTGTGCTTCTTGAAGGGGACATTGTATATGT	1496
QY	480	EGLUGLYALATHRGILUGLULYBLEULEUSERARGLYSHECTELYRTRYTRPALATH	500
DB	1497	CGAAGAGCCACAGGAGGAGGAGATTAAGTACGCGGAGAAATGATGAATAATCTGGGCTAC	1556
QY	500	RPHEHARGTHRGLYANPROMANGLYASNAASPLEUSERLEUTRPPROVALTYRANLE	520
DB	1557	CTTGTGCTGAACCGGAAATCTTAATGGGAAAGCACTGTCTCTGTGGCCAGCTTATATCT	1616
QY	520	UTHRGILUNTRYLEUGLINLEUASPLEUASMETSERLEUGLYINARGLEULYSGILUPR	540
DB	1617	GACTGAGAGTACCTGCAAGCTGGAATTGACATAGAGCTCGGAGACGACATCAAGAAACC	1676
QY	540	CAARGVALASPHETRTPHRSERTHRIIEPROLEUILLLEUSERALASERASPMETLEUHI	560
DB	1677	GCGGCTGAGTGTGGACACACACATCCCTGATCTGTGCTGCTCGACATGCTCCA	1736

Qy	560	bSERProLeuSeSerLeuThrPheLeuSerLeuLeuGlnGlnProPhePhePheCysAl	580
Dd	1737	CAGGCCCTCTTCTCCTTAACCTTCCCTCTCTCCCAAGCCTTCTTTCTTTTGAGC	1796
Qy	580	APRO 581	
Dd	1797	TCCCT 1800	
RESULT 7			
ID	AB086170	standard; DNA; 1746 BP.	
AC	AB086170;		
XX			
DT	10-SEP-2002	(first entry)	
XX			
DE	Novel human gene. SEQ ID 41.		
XX			
KW	Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;		
KW	neurotropic; neuroprotective; immunosuppressive; haemostatic;		
KW	antiinflammatory; cardiac; antitumor; vitruide; antithyroid;		
KW	cardioprotective; anorectic; metabolic; vaccine; cancer; infection;		
KW	wound healing disorders; atherosclerosis; Parkinson's disease;		
KW	Alzheimer's disease; autoimmune disorder; haematopoietic disorder;		
KW	inflammation; neoplastic disease; nervous system disorder;		
KW	cardiovascular disorders; pancreatitis; respiratory disorder;		
KW	hyperproliferation; systemic autoimmune disease; hyper-immunity;		
KW	developmental abnormality; gastrointestinal ulceration; neuropathy;		
KW	haematological disease; metabolic disease; sperm dysfunction;		
KW	thyroid disorder; hypothyroidism; brain damage; colitis;		
KW	cone photo- transduction deficiency; neurological disease; stroke;		
KW	angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;		
KW	trachea; thymus; lymph node; muscular system; obesity; anorexia;		
KW	growth abnormality; precocious puberty; gene; ss.		
OS	Homo sapiens.		
XX			
PN	WO200250105-A1.		
XX			
PD	27-JUN-2002.		
XX			
PF	17-DEC-2001; 2001WO-US049232.		
XX			
PR	19-DEC-2000; 2000US-0256710P.		
PR	20-DEC-2000; 2000US-0257048P.		
PR	09-JAN-2001; 2001US-0260482P.		
PR	30-JAN-2001; 2001US-0264922P.		
PR	06-FEB-2001; 2001US-0266797P.		
PR	19-MAR-2001; 2001US-0276986P.		
PR	04-APR-2001; 2001US-0281535P.		
PR	08-MAY-2001; 2001US-0289622P.		
XX			
PA	(SMIK) SMITHKLINE BEECHAM CORP.		
PA	(SMIK) SMITHKLINE BEECHAM PLC.		
PA	(GLAX) GLAXO GROUP LTD.		
XX			
PI	Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;		
PI	Mattensen SA, Ritzvi SK, Smith RF, Strum UC, Xie Q;		
XX	WPI; 2002-508784/54.		
DR	P-PSDB; ABP61005.		
XX			
PT	Secreted proteins and polynucleotides useful as vaccines for preventing		
PT	or treating various diseases e.g. cancer, wounds, atherosclerosis,		
PT	Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.		
XX			
PS	Claim 2(a); Page 251-252; 335pp; English.		
CC	The invention relates to an isolated polypeptide with signal sequences		
CC	which allow it to be secreted extracellularly or membrane associated. The		
CC	activity of polypeptides of the invention may be described as,		
CC	cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic,		
CC	neuroprotective, immunosuppressive, haemostatic, antiinflammatory,		

[illegible]

PR 18-JUN-2001; 2001US-0298952P.
 PR 19-JUN-2001; 2001US-0299237P.
 PR 19-JUN-2001; 2001US-0299276P.
 PR 12-SEP-2001; 2001US-0318750P.
 PR 25-SEP-2001; 2001US-0324800P.
 PR 25-SEP-2001; 2001US-0324802P.
 PR 17-OCT-2001; 2001US-0325684P.
 PR 14-NOV-2001; 2001US-0330143P.
 PR 14-NOV-2001; 2001US-0332131P.
 PR 14-NOV-2001; 2001US-0332240P.
 PR 14-NOV-2001; 2001US-0332779P.
 PR 21-NOV-2001; 2001US-0332115P.
 PR 04-DEC-2001; 2001US-0337621P.
 PR 03-JAN-2002; 2002US-0345783P.
 PR 16-JAN-2002; 2002US-0350251P.
 PR 02-APR-2002; 2002US-00114270.
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 PI Guo X, Kekuda R, Miller CE, Malysankar UM, Spylek KA;
 PI Paturajan M, Liu X, Gusev VY, Li L, Vermet CAM, Zehusuen BD;
 PI Gorman L, Shenoy SC, Pena CRA, Smitheon G, Burgess CE, Gerlach V;
 PI Padigara M, Shinkets RA, Gangoli BA, Taupier RJ, Caeman SJ, Ji W;
 PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ;
 PI Macdonald JR, Rothenberg MB, Mazur A, Millet I, Peyman JA;
 PI Ellerman K;
 XX
 XX WPI; 2003-046858/04.
 DR P-PSDB; ABUS4639.
 XX
 PT New isolated NOVX polypeptide useful for treating atherosclerosis,
 PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
 PT neurodegenerative disorders, Alzheimer's disease and cancer.
 XX
 PS Claim 17; Page 303; 666BP; English.
 XX
 CC The invention relates to human polypeptides, termed NOVX, and the
 CC polynucleotides encoding them. The polypeptides and polynucleotides are
 CC useful for diagnosing disease, and screening for potential therapeutic
 CC agents. The sequences are useful for treating metabolic disorders,
 CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
 CC stenosis, atrial septal defect (ASD), aortic ventricular canal defect,
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
 CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
 CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
 CC and cancer. Sequences ABX72170-ABX72275 represent human NOVX
 CC polynucleotides of the invention
 XX
 SO Sequence 1746 BP, 402 A; 457 C; 459 G; 428 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.93e-236 Length: 1746
 Score: 2597.50 Matches: 497
 Percent Similarity: 95.07% Conservative: 4
 Best Local Similarity: 94.31% Mismatches: 17
 Query Match: 84.36% Indels: 9
 DB: Gaps: 3
 US-10-023-515-2 (1-581) x ABX72267 (1-1746)
 QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrPileGlnGlyLeuGln 50
 DB 79 GGGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGGCAAGCAA 138
 QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIa 70
 DB 139 GTCACTGTCGCGGAGAGCCCTGCTGTGAACGCTTCTCTCGAGAGTCCCTTCTCTGCT 198
 QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeu 90
 DB 199 CCGCGGCTGGGATCCCTCGATTTCAGAACCGGCAAGCCGTGATGCGCTGGGATATACCTTG 258

QY 91 ArgGluAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGluTyrPheLeuLeuAsp 110
 DB 259 CGAGAGGCACTCTCTACCTTAATTTGCTCCAGAACTCAGAGTGGCTGCTTAAT 318
 QY 111 GlnHisMetLeuLeuValHisTyrProValPheGlyValSerGluAspCysLeuTyrLeu 130
 DB 319 CAACATATGCTCAAGTGCATTACCCGAAATTCGGAGGTGCAAGAACTGCTTACCTG 378
 QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerIleuProValLeuValTyrPhe 150
 DB 379 AACATCTATGCGCTGCCACCGCATACAGGCTCCAGGCTCCCGCTTGCTGCTTC 438
 QY 151 ProGlyGlyAlaPheLeuTyrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa 170
 DB 439 CCAAGAGTGCCTTCAAGACTGCTCAGCTCCATCTTGAATGGTCCGCTGCTGCTCC 498
 QY 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
 DB 499 TATAGAGACGTGCTGTTGTGTGCTCAGTACCGGCTAAGAAATTTGGTTCTTCAAC 558
 QY 191 ThrTyrAspGlnHisAlaProGlyAsnTyrAlaPheValAspGlnValAlaAlaLeuSer 210
 DB 559 ACATGGGATCAGCATGCTCCGGGGAACGGGCTTCAGAGAACAGGTGGCTGCTGCTC 618
 QY 211 TrpValGlnValAsnIleGluPhePheGlyValAspProSerSerValThrIlePheGly 230
 DB 619 TGGCTCCAGAGAACATCAGATGCTTCTCGTGGGGAACCCAGCTCTGTACATCTTTGGC 678
 QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIleGlyLeu 250
 DB 679 GAGTCCGGGGAGGCAATAGTGTTCATGCTTATAGTGTCTCCATGCGCAAGGCTTA 738
 QY 251 PheHisIleValAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAsp 270
 DB 739 TTCCACAAAGCCATCATGAGAGTGGGCTGCGCATCATCTTCACTTGAGAGCCCATAT 798
 QY 271 TyrGluValSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAsp 290
 DB 799 TATGAGAGAGTGAAGACCTCAGGTGTGTCACTTCTGTGTAACATGCGTCAAC 858
 QY 291 SerGluAlaLeuLeuArgCysLeuArgThrValProSerIleGlyLeuLeuThrLeuSer 310
 DB 859 TCTGAGGCGCTCGTGAAGTGTCTGAGGACAAACCTTCAGAGAGCTGACCTGACGC 918
 QY 311 GlnIleThrIleSerPheThrArgValValAspGlyValaPhePheProAsnGluProLeu 330
 DB 919 CAGAAACAAAGTCTTTCACCTCGAGTGTGATGTGCTTCTTCTTCAATGAGGCTCTTA 978
 QY 331 AspLeuLeuSerGlnValAlaPheValAlaIleProSerIleIleGlyValAsnAsnHis 350
 DB 979 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCAATCATGAGTCAATAACAC 1038
 QY 351 GluCysGlyPheLeuLeuProMetLeuGlu-----AlaProGluIleLeuSerGlySer 368
 DB 1039 GAGTGTGCTTCTCGTCTGCTATGTAAGATTCGTGGCTGTCACATCCACTCCCTCA 1098
 QY 369 AsnIleSerIleValLeu-----HisLeuIleGlnAsnIleLeuHisIle 383
 DB 1099 AACCGTATGCAAGCTTGGCTTCAACAGCTGGGCAATTCACAGAGAA-----CATATC 1152
 QY 384 ProProGlnTyrIleHisIleValAlaAsnGluTyrPheHisAspIleHisSerLeuThr 403
 DB 1153 CCGGCTCGATTTGTCACCTTGTGCTATATAATCTTCATGACAGACACTCCCTGACT 1212
 QY 404 GluIleArgAspSerLeuLeuAspLeuGlyValPhePheValValProAlaLeu 423
 DB 1213 GAATTCGAGACAGTCTTCTGCACTGCTTGAATGATGTTCTTTGTGTGCTCCCTGACTG 1272
 QY 424 IleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHis 443
 DB 1273 ATCAAGCTCGATATATCACAGAGATGCTGTGCACTGTCTTATGAGATTTGCGGAC 1332


```

Oy 231 GUSerAlaGlyAlaIleSerValSerLeuIleLeuSerProMetAlaIleGlyLeu 250
Db 874 GAGTCCGGGGAGCCATAAGTCTTCTAGCTTATACGTCTCCCAATGGCCAAAGCTTA 933
Oy 251 PheHisIlyAlaIleMetGlySerGlyValAlaIleIleProTyLeuGluAlaHisAap 270
Db 934 TTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGGCCCATGAT 993
Oy 271 TyTGlyIlySerSerGlyAapLeuGlnValAlaHisIlePheCysGlyAapAapAaSerAap 290
Db 994 TATGAGAAAGAGAGAGACCTGAGAGTGGTGGTGCATTTCTGTGTGAACAAATGGCTCAGAC 1053
Oy 291 SerGluAlaLeuLeuArgCysLeuArgThrIlySerProSerIlyGlyLeuLeuThrLeuSer 310
Db 1054 TCTGAGGCCCTCTGCTAGAGTGGCTGAGACAAACCTCCAAAGAGACTGCTGACCTCAGC 1113
Oy 311 GlnIlyThrIlySerPheThrArgValAlaAapGlyAlaPhePheProAapGlnProLeu 330
Db 1114 CAGAAAACAAAGTCTTTCATCTGAGTGGTGTGATGGTCTTCTTCTTATATGAGCTCTTA 1173
Oy 331 AapLeuLeuSerGlnIlyAlaPheIlyAlaIleProSerIleIleGlyValAapAapHis 350
Db 1174 GATCATATTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGTCAATTAACAC 1233
Oy 351 GluCysGlyPheLeuLeuProMetIlyGlyAlaProGlnIleLeuSerGlySerAapIly 370
Db 1234 GAGTGTGGCTTCTCTCTGCTCTATG----- 1257
Oy 371 SerLeuAlaLeuHisIleLeuIleGlnAapIleLeuHisIleProProGlnIlyThrLeuHisIle 390
Db 1258 -----CACATCCGCCCTCAGATATTGACACTT 1284
Oy 391 ValAlaAapGlnIlyThrPheHisAapIlyHisSerLeuThrGlnIlyAapSerLeuLeu 410
Db 1285 GTGGGTATATGATATCTTCCATGACAAAGCACTCCGACTGGAATCCGAGACAGTCTTCTG 1344
Oy 411 AapLeuLeuGlyAapValPhePheValAlaProAlaLeuIleThrAlaArgTyHisAap 430
Db 1345 GACTTGTCTGAGAGTGTCTTCTTGTGTCTCTGACATGATACAGCTCGATATACAGA 1404
Oy 431 AapAlaGlyAlaProValTyThrPheTyGlnPheArgHisAapProGlnCysPheGlnAap 450
Db 1405 GATGCTGGTGACCTGCTACTTCTATGAGTTGGGACCGGCTCAGTGGCTTTGAAAGAC 1464
Oy 451 ThrIlyPProAlaPheValIlyAlaAapHisAalAapGlyValAapPheGlyGly 470
Db 1465 ACGAAGCCAGCTTTGTCAAAGCCAGCAACGCTGATGAAGTCCGCTTGTGTCCGTGGT 1524
Oy 471 AlaPheLeuIlyGlyAapIleValMetPheGlnGlyAlaThrGlnGlnIlyLeuLeu 490
Db 1525 GCCTTCTCGAAGGGGACATGTGTATGTTCGAAGAGCCACGAGAGAGAGAGATTAAGT 1584
Oy 491 SerArgIlyMetMetLeuTyThrPalaThrPheAlaArgThrGlyAapProAapGlnIlyAap 510
Db 1585 AGCCGGAAGATATGAATACTAGGCTACCTTGTCTGAAACCGGAAATCTTAATGGAAAC 1644
Oy 511 AapLeuSerLeuTyThrProAlaTyThrAapLeuThrGlnIlyLeuGlnIlyAapLeuAap 530
Db 1645 GACCTGTCTGTGTGGCCAGCTTATATCTGACTGAGAGTACCTCAGGTGAGACTTTGAAC 1704
Oy 531 MetSerLeuGlyGlnAapLeuIlyGlyProAapGlyValAapPheThrPheSerThrIlePro 550
Db 1705 ATGAGCCCTCGACAGAGACTCAAAAGAACCGGAGAGAGATGTGGGTGACGGGGTATCTCT 1764

```

```

XX cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.
KW
XX
XX Felis catus.
OS
XX JP2003250575-A.
PN
XX
XX
PD 09-SEP-2003.
XX
XX 04-MAR-2002; 2002JP-00057908.
XX
XX 04-MAR-2002; 2002JP-00057908.
XX
XX (TOHO-) TOHOKU TECHNOARCH KK.
PA
XX
XX WPI, 2004-002277/01.
DR
XX P-PSDB; ADF50147.
XX
XX Novel cauxin protein or its salt, useful as a cat kidney disease marker,
PT and for diagnosing cat kidney disease.
XX
XX Claim 6; SEQ ID NO 3; 33pp; Japanese.
XX
XX The present sequence encodes a cat cauxin protein (I) or its salt, which
CC is cat kidney disease marker. Also described: (1) a partial peptide (II)
CC of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
CC comprising (III); (4) a transformed host (V) comprising (III) or (IV);
CC (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
CC couples specifically with (I) or (II); (7) diagnosing cat kidney disease
CC which involves measuring (I) quantitatively, and where reduction of
CC amount of (I) indicates presence of the disease; (8) a cat kidney disease
CC diagnostic agent comprising (I) labelling agent, a reagent which measures
CC the biological activity of urinary (I) or (VI); and (9) a cauxin
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.
XX
XX SQ Sequence 2145 BP; 444 A; 581 C; 550 G; 564 T; 0 U; 6 Other;
XX
XX Alignment Scores:
XX Pred. No.: 5,74e-181 Length: 2145
XX Score: 2015.00 Matches: 388
XX Percent Similarity: 80.87% Conservative: 60
XX Best Local Similarity: 70.04% Mismatches: 102
XX Query Match: 65.44% Indels: 4
XX DB: Gaps: 2
XX
XX US-10-023-515-2 (1-581) x ADF50146 (1-2145)
Oy 31 GlyProSerAlaGlyAlaIleSerValSerLeuIleLeuSerProMetAlaIleGlyLeu 50
Db 228 GGGCCAGCTGTGAGTGAACAGAGTGAAGAGCAACAGCTGGAGTGGTCCGGGGAAAGCAA 297
Oy 51 ValThrValIleGlySerProValAapLeuValPheLeuGlyValAapPheHisAalAa 70
Db 228 ACCATGTACTGGGAAGACCGCTGCTGAAACAGATTCCTCGGAGTCCCTATCTCTGCA 357
Oy 71 ProProLeuGlySerLeuArgPheThrAapProGlnProAlaSerProTyPheAapAap 90
Db 358 CTTCTCTAGAGCCCTCTGCAATTTAAGCAACAAAGCTGCTGCTCCGGAAATGACTTC 417
Oy 91 ArgGluAlaThrSerTyThrProAapLeuCysLeuGlnAapSerGlyTyThrLeuLeuAap 110
Db 418 CGAAATGCAACATCTTACCTAAATTAATGCTTCCAGACTTAAGATGGCTGCTCTAT 477
Oy 111 GlnHisMetLeuIlyValHisIlyTyThrProIlyPheGlyValSerGlyAapCysLeuTyThr 130
Db 478 CAACAGCTTCTCAAGTCCAGTATCCCAATTTGAAGGCTCGAAGAGCTGCTGTAACCTT 537
Oy 131 AapIlyTyAlaProAlaHisAalAapThrGlySerTyIlyLeuProValIleValIlyPhe 150

```

```

|||||
538 AACATCTATGCGCCAGCCCATGCGGACATATGCTCCAACTCTCTGTATGATGCTGTC 597
|||
QY 151 ProGlyGlyAlaPheLeuTherGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
|||
Db 598 CCGGGGGGGCTTCTCAAGATGGGCTCAGCTTCTCTCGATGGGTCCGCTTGCGCTGCC 657
|||
QY 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
|||
Db 658 TACGAGGAGCGTGTGATGTGATCCAGTACCGGCTAGGAAATATGTTTGTTCGAC 717
|||
QY 191 ThrTPAPAPGlnIleAlaProGlyAsnTPAlaPheLeuAspGlnValAlaAlaLeuSer 210
|||
Db 718 ACAGGGGATGAGCAATGCCCCGGGGAACTGGGCTTGGCGACAGCTGGCGCTCAAC 777
|||
QY 211 TrpValGlnIleValAsnIleGluPhePheGlyGlyValAspProSerSerValThrIlePheGly 230
|||
Db 778 TGGGTCCGGGACAAACATGAGATCTTTCGGTGTGACCAAGCTCCGACCATCTTTTGA 837
|||
QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 250
|||
Db 838 GAGTCACGGGAGCCATCAGTGTTCAGCTCTATCTGTCCCAATGACCAATGCTTA 897
|||
QY 251 PheIleTyrAlaIleMetGluSerGlyValAlaIleIleProTyrLeu--GluAlaHis 269
|||
Db 898 TTCACAAAGCCCATCAGAGAGTGGGGTGCATCTGCTTACTGATGAGACCCCT 957
|||
QY 270 AspTyrGluLysSerGluAspLeuGlnValAlaAlaIlePheCysGlyAsnAsnIleSer 289
|||
Db 958 GGTGATGAGAGAAAGATTTGCAAGGTCTTCCCGCTGCTGTGTGTTCCATGGGCT 1017
|||
QY 290 AspSerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeu 309
|||
Db 1018 GACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
|||
QY 310 SerGlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluPro 329
|||
Db 1078 AGCAAGAACTACACGTTTCCATCCAGATTTGATGATCTTTTCTTCTCGATAGGCT 1137
|||
QY 330 LeuAspLeuLeuSerGlnLysAlaPheLeuValaIleProSerIleIleGlyValAlaAsn 349
|||
Db 1138 GTAGCCCTATTATCACTCAAAACATTTAATTCAGTCTCTTATCATCGAGTCAATAC 1197
|||
QY 350 HisGluCysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsn 369
|||
Db 1198 CACGAGTGTGCTTCTCTG---TCCACGAGTGTCTGAGATCTCTGGGGGCTCCAAAC 1254
|||
QY 370 LysSerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHis 389
|||
Db 1255 AGGTCTGCGCCCTTACTTACTAGTACACAGTCTCCGAAATTTCCACACCATTTTGCAC 1314
|||
QY 390 LeuValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleLeuAspSerLeu 409
|||
Db 1315 CTTGGGTGATGATTTACTTCTTACAAACAGACTCCCGGTGAAATACAGATAGTTT 1374
|||
QY 410 LeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHis 429
|||
Db 1375 CTGACCTTGCTTGAAGATGCTCTTGTGCTCTGCTGGGGTGTGACAGCTCATATCAT 1434
|||
QY 430 ArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnIleCysPheGlu 449
|||
Db 1435 AGAATATCTGTGACCTGTCTTACTTCTATGAGTTTAAACACCCGCCAGTCTTAAAC 1494
|||
QY 450 AspThrLysProAlaPheValLysValAspHisAlaAspGluValArgPheValPheGly 469
|||
Db 1495 GACACGAGGCGCAGCTTCTGTGAAGCGCATCTGATGAATCCGCTTCGCTTGTGA 1554
|||
QY 470 GlyAlaPheLeuLysGlyAspIleValMetPheGluGlyValAlaThrGluGluLysLeu 489
|||
Db 1555 GGTGCTTCTTGAAGCGCATGTCATGTTGAAAGAGGCCACGAGAGGAGAAATTG 1614
|||
QY 490 LeuSerArgLysMetLeuTyrTyrPAlaThrPheAlaArgThrLysAsnProAsnGly 509
|||

```

```

Db 1615 CTGAGCAGGAAATGATGAGTACTGCGCAACTTTGCTCGGACCGGGGACCTTAACGGG 1674
QY 510 AsnAspLeuSerLeuTyrProAlaTyrAsnLeuThrGluGlnIleLeuAspLeu 529
|||
Db 1675 GAAAGTGTCTCTTGTGGCGCACCTTACACCCAGACGAGACGATACCTGAAGCTGATTTG 1734
|||
QY 530 AsnMetSerLeuGlyGlnArgLeuLysGluProArgValAspPheThrIle 549
|||
Db 1735 AGTGTAGCGCTGTGACAGAACTGAAGAGCAAGAGTGTGTGATGAATACATTT 1794
|||
QY 550 --ProLeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPhe 569
|||
Db 1795 GTCCCTGATACCCCCCACTTCAGGGCCCTCCCAAGTCTCTTTCCTTACTCTCC 1854
|||
QY 569 euSerLeuLeuGlnProPhePhePheCysAlaPro 581
|||
Db 1855 TTCCTTGTCTCCCGCTTGTCTTCTTCTGCTCA 1892
|||
RESULT 12
ADP50144
ID ADP50144 standard; cDNA; 1629 BP.
XX
AC ADP50144;
XX
DT 12-FEB-2004 (first entry)
XX
DE Cat cauxin encoding cDNA SEQ ID NO:1.
XX
KW cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.
XX
OS Fells catus.
XX
PN JP2003250575-A.
XX
PD 09-SEP-2003.
XX
PE 04-MAR-2002; 2002JP-00057908.
XX
PR 04-MAR-2002; 2002JP-00057908.
XX
PA (TOHO-) TOHOKU TECHNORCH KK.
XX
DR WPI; 2004-002277/01.
XX
DR P-PSDB; ADP50145.
XX
PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX and for diagnosing cat kidney disease.
XX
PS Claim 6; SEQ ID NO 1; 33p; Japanese.
XX
XX
The present sequence encodes a cat cauxin protein (I) or its salt, which
is cat kidney disease marker. Also described: (I) a partial peptide (II)
of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
comprising (III); (4) a transformed host (V) comprising (III) or (IV);
(5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
couple specifically with (I) or (II); (7) diagnosing cat kidney disease
which involves measuring (I) quantitatively, and where reduction of
CC amount of (I) indicates presence of the disease; and
CC diagnostic agent comprising (I) labelling agent; (8) a cat kidney disease
CC the biological activity of urinary (I) or (VI); and (9) a cauxin
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.
XX
XX
Sequence 1629 BP; 345 A; 438 C; 441 G; 405 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,73e-178 Length: 1629
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60

```

```

Best Local Similarity: 71.73% Mismatches: 85
Query Match: 64.47% Indels: 2
DB: 12 Gaps: 2
US-10-023-515-2 (1-581) x ADF50144 (1-1629)

QY 31 G1PProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnIleGlySerGln 50
DB 64 GGGCCAGCTGCTGATGACACACAGAGAGACACAGGCTGGATGGGTCCGGGGAGAACAA 123
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 124 ACCACTGACTAGGGAGACACCTGGCTGTGAACATGTTCTCGGGATCCCTATGCTGCA 183
QY 71 ProPheLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAsnLeu 90
DB 184 CCTCTCTAGGGCCCTCGGATTTAAACAACAAAGCCGTCTGCGCCGGAAATGACTTC 243
QY 91 ArgGlnAlaThrSerTyrrProAsnLeuCysLeuGlnAsnSerGluTrpLeuLeuAsp 110
DB 244 CGAATGCCACATCTCTACCTTAATTAATGCTTCCAGGACTTAGAGTGGCTGCTCTTAT 303
QY 111 GlnHisMetLeuValHisIleTyrrProIysPheGlyValSerGluAspCysLeuTyrrLeu 130
DB 304 CAACAGTTCTCAAGTGGGTTAACCCAAATTGGAAGCTCCGAAGACTGCTGTACCTT 363
QY 131 AsnIleTyrrAlaProAlaHisAlaAspThrGlySerIleValLeuValTrpPhe 150
DB 364 AACATCTTAGCGCCAGCCCATGCGACAAATGGCTCCACCTCCGTCAATGGTGTGCTTC 423
QY 151 ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
DB 424 CCGGGGGGTGCTTCAAGATGGGCTCAAGTCTCTCTCTGATGGGTCCGCTTGGCTGCC 483
QY 171 TyrrGluAspValLeuValValValGlnTyrrArgLeuGlyIlePheGlyPhePheThr 190
DB 484 TACAGAGAGCTGCTGATGATGATACCCAGTACCGGCTGAGAAATTTGGTTTTCGAC 543
QY 191 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheIleAspGlnValAlaAlaLeuSer 210
DB 544 ACAGGGAGATGAGCAATGCCCGGGGAACTGGGCTTGTGGAGACAGTGGCTCCCTCAAC 603
QY 211 TrpValGlnIysAsnIleGluPhePheGlyIleAspProSerSerValThrIlePheGly 230
DB 604 TGGGTCGGGACACATGATGATTTCTTGGTGTGACCAACGCTCCGTGACATCTTTGAA 663
QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIysGlyLeu 250
DB 664 GAGTACAGCGGAGCATCATGATTTCCAGCTCATTTCTGTCCTCCCATGACCAATGGCTTA 723
QY 251 PheHisValAlaIleMetGluSerGlyValAlaIleIleProTyrrLeu---GluAlaHis 269
DB 724 TTCACAAAGCCATCATGAGAGTGGGGGCGCATCTGCTTTACTGATGAGACCCCT 783
QY 270 AspTyrrGluIysSerGluAspLeuGlnValAlaAlaHisPheCysGlyAsnAsnAlaSer 289
DB 784 GGTATGAGAGGAGAGAGATTTGCGAGTGTGGCCGATTTGTGGTTGCCATGGGTCT 843
QY 290 AspSerGlnAlaLeuLeuArgCysLeuArgThrIysProSerIysGlyLeuLeuThrLeu 309
DB 844 GACTCTGCTGCTCTGCTGAGTGTGCGAGGCAAAACCTCGAGAGAGTGTATGACATC 903
QY 310 SerGlnIysThrIysSerPheThrArgValValAspGlyAlaPhePheProAsnGluPro 329
DB 904 AGCAAGAACTCACCTTTTCCATTCAGATGATGATCTTTCTTCTTCATGATGAGCT 963
QY 330 LeuAspLeuLeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAsnAsn 349
DB 964 GTAGCCCTATTGACTCAAAAGCATTTAATTCATGCTTCTTATCATCGAGATCAATAC 1023
QY 350 HisGluCysGlyPheLeuLeuProMetIysGlnAlaProGluIleLeuSerGlySerAsn 369
DB 1024 CACGAGTGTGCTCTCTCTCTG---TCCACGAGATTTTTCGATGATCTCGGGGGCTTCAAC 1080

```

```

QY 370 LysSerLeuAlaLeuHisIleLeuIleGlnAsnIleLeuHisIleProProGlnIleTyrrLeuHis 389
DB 1081 AGCTCTTGGCCCTCTTACTTACTTACACACGCTTCTGAAATATTTCCACCAAGATTTTGGAC 1140
QY 390 LeuValAlaAsnGluTyrrPheHisAspIysHisSerLeuThrGluIleArgAspSerLeu 409
DB 1141 CTTTGGGCTGATCATTTACTTCTACAAACAGCACTCCCGCGTTGAAATATAGATATGATTTT 1200
QY 410 LeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrrHis 429
DB 1201 CTGACCTTGGCTTGAGATGTGCTTTTGTGTCTTGGGCTGTGTACACCTGATATTCAT 1260
QY 430 ArgAspAlaGlyAlaProValTyrrPheTyrrGluPheArgHisArgProGlnCysPheGlu 449
DB 1261 AGAGATGCTGGTGCACCTGTCTTCTTATGATTTCAACACCGCCGCAAGTGTCTTAAC 1320
QY 450 AspThrIysProAlaPheValIysAlaAspHisAlaAspGluValArgPheValPheGly 469
DB 1321 GACACGAGGCCAGCTTTGTGAAAGCCCATCATCTGATGAAATCCGCTTGTGGA 1380
QY 470 G1YAlaPheLeuIysGlyAspIleValIleMetPheGluGlyAlaThrGluGluIysLeu 489
DB 1381 GGTGCTTCTCTGAAAGGCGACATTTGTATGTTTGAAGAGGCCACCGAGAGAGAAATTG 1440
QY 490 LeuSerArgIysMetMetIysTyrrTrpAlaThrPheAlaArgThrGlyAsnProAsnGly 509
DB 1441 CTGACAGAGAAATATATAGATGATGAGGCACTTGTGCTGCGAACCGGGACCTTAACGGG 1500
QY 510 AsnAspLeuSerLeuTrpProAlaTyrrAsnLeuThrGluGlnIleLeuAspLeu 529
DB 1501 GAAGGTGTGCTCTTGTGGCGACCTTACACCCAGAGCGAGCATCTGTAAGCTGATTTG 1560
QY 530 AsnMetSerLeuGlyGlnArgIleuIysGluProArgValAspPheThrTrpSerThrIle 549
DB 1561 AGTGTGACGCTGTGGACAGAACTGAAGAGCAAGAGTGAGATTGTGATGATATTCATT 1620

RESULT 13
AB086171
ID AB086171 standard; DNA; 1071 BP.
XX
AC AB086171;
XX
DT 10-SEP-2002 (first entry)
XX
DE Novel human gene. SBO ID 42.
XX
KW Human; cytosolic; vulnary; antiarteriosclerotic; antiparkinsonian;
KW neurotrophic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antilicer; virucide; antithyroid;
KW cerbroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haemotopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW angio genesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200250105-A1.
XX
PD 27-JUN-2002.
XX
PF 17-DEC-2001; 2001WO-US049232.
XX
PR 19-DEC-2000; 2000US-0256710P.

```

PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264923P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,
PI Mathensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX WPI; 2002-508784/54.
DR P-PSDB; ABP61006.
XX
PT Secreted proeline and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 2(a); Page 252; 335pp; English.
XX
CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosolic, vulnary, antiarteriosclerotic, antiparkinsonian, nootropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, anticancer, virucide, antihypoid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder.
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, anglogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records AB086130-AB086164 represent novel human cDNA's
CC of the invention
XX
SQ Sequence 1071 BP; 242 A; 286 C; 254 G; 289 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4,46e-159 Length: 1071
Score: 1780.00 Matches: 340
Best Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.81% Indels: 0
DB: 6 Gaps: 0
US-10-023-515-2 (1-581) x AB086171 (1-1071)
QY 242 I l e u s e r p r o m e t a l a y s g l y l e u p h e i s u y a l a i l e m e t g l u s e r g l y v a l a l a 261
DB 49 A T A C T G C T C C C A A G G C C A A A G G C T A T T C C A A A G C A C A T C A T G A G A G G G G T G C C 108
QY 262 I l e l l e p r o t y r l e u g l u a l a h i a p r t y r g l u y s e r g l u a p l e u g l n v a l a l a 281
DB 109 A T C A T C C C T T A C C T G A G G C C C A T A T T A T A G A A G A C T G A G A C C T G C A G T G T G C A 168
QY 282 H i s p h e C y s g l y a n a a n u l a s e r a s p s e r g l u a l e u a u a g C y l e u a r t h r i y s 301
DB 169 C A T T T C T G T G A C A A T G C C T C A G A C T C T A G G C C C T G A G G T G C T T A G A G A C A A A 228
QY 302 P r o s e r l y s g l u e u l e u t h r l e u s e r g l n y s t h r l y s e r p h e t h a r v a l a p 321
|||||

DB 229 C C C T C C A A G A G A C T G C T G A C C C T C A G C C A G A A A C A A A G T C T T T C A C T G A G T G T G A T 288
QY 322 G l y l a p h e p h e p r o a n g l u p r o l e u a s p r l e u s e r g l n y s a l a p h e l y a l a l l e 341
DB 289 G A T G C T T C T T C C A A T G A G C C T C T G A T C T A T T G T C T G A A A G C A T T T A A A G C A A T T 348
QY 342 P r o s e r l l e i l e g l y a l a s n a s n h i s g l u C y s g l y p h e l e u l e u p r o m e t l y s g l u a l a 361
DB 349 C C T T C A T C A T C G A G T C A A T A A C A G A G T G G C T T C C T G C T G C C A T G A T A A G A G A G C T 408
QY 362 P r o g l u l l e u s e r g l y s e r a s n l y s e r l e u a l a e u h i s l e u i l l e g l a s n l l e u 381
DB 409 C C T G A G A T C C T C A G T G G C T C C A A C A G C C C T T G C C C C T C C A T C G A T C A A A A C A T C C T G 468
QY 382 H i s l l e p r o p r o g l n t y r l e u h i s l e u v a l a a n g l u t y r p h e i s a s p l y s e r 401
DB 469 C A C A T C C G C C T C A G T A T T T G C A C C T T G T G C T A T G A T A T C T T C A T G A C A A G A C A C C C 528
QY 402 L e u t h r g l u l l e a r a s p s e r l e u a s p r l e u a s p r l e u g l y a a p v a l p h e p h e v a l p r o 421
DB 529 C T G A C T G A A A T C G A G A C A G C T T C T G A C T T G C T T G S A G A T G T T C T T T G T G T C C C T 588
QY 422 A l a l e u l l e t h r a l a r g t y r h i a r g a s p l a g l y a l a p r o v a l t y r p h e t y r g l u p h e 441
DB 589 G C A C T G A T C A C A G C T C A T A T C A C A G A G A T G C T G T G C A C C T G T C T A T G A G A G T T 648
QY 442 A r g h i s a r p r o g l n y s p h e g l u a s p r t h r l y p r o a l p h e v a l l y a l a s p h i a l a 461
DB 649 C G G A C C G G C T C A G T G C T T T G A A G A C A G A A G C A G T T T G T C A A A G C C G A C C A C C T 708
QY 462 A s p g l u a l a r p h e v a l p h e g l y l a l a p h e l e u y s g l y a a p l l e v a l m e t p h e g l u 481
DB 709 G A T A A G T C G C C T T G T G T T G G T G G T G C C T T C T G A A G G G G A C A T T G T A T G T T G A A 768
QY 482 G l y a l a t h r g l u g l u l y l e u l e u s e r a r g l y s e t e c t l y s t y r t r p a l a t h r p h e 501
DB 769 G G A C C A C G G A G G A G A A G T T A C T A G C C G A A G A T G A A A T C T G G G C T A C C T T 828
QY 502 A l a a r g t h r g l y a n p r o a n g l y a s n a s p r l e u t r p r o a l t y r a s n l e u t h r 521
DB 829 G C T G A A C C G G A A T C C T A T G S A A G A C A C T G T C T G T G C C A G C T T A T A T C T G A C T 888
QY 522 G l u l i n t y r l e u g l n l e u a s p r l e u a s m e t s e r l e u g l n a r g l e u l y p r o a r g 541
DB 889 G A G A G A T A C C T C C A G C T G A C T T G A C A T G A C G C T C C G A C A G A C T C A A A G A C C G G C G 948
QY 542 V a l a s p h e t r p h r s e r t h r l e p r o l e u l l e u s e r a l a s e r a s p m e t l e u h i s e r 561
DB 949 G T G A T T T T G A C A C A C A C A C A T C C C C T G A T C C T G C C T C C G A C A T G C T C A C A G T 1008
QY 562 P r o l e u s e r l e u t h r p h e l e u s e r l e u g l n p r o p h e p h e p h e C y a l a p r o 581
DB 1009 C C T C T T C T C T T A A C T T C T C T C T C C T C C A G C C T T C T T T T T T G T G T C C T 1068
RESULT 14
ADCS5523 ID ADCS5523 standard; cDNA; 1244 BP.
XX
AC ADCS5523;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human carboxylatase 24.64 encoding sequence.
XX
KW human carboxylatase-24.64; primary hypertension; digestive ulcer;
KW nephrotic; bronchial asthma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 236..910
FT /*tag= a
XX

PS Disclosure; Fig 8A; 53bp; French.

XX The present sequence represents the coding region of human intestinal
 CC carboxylesterase (ICE) gene. ICE induces specific cytotoxic T
 CC lymphocytes (CTL) and secretion by these CTL of cytotoxic factors, e.g.
 CC interleukin-2, interferon gamma and tumour necrosis factor. ICE
 CC polypeptides and polynucleotides are used for treating cancer, by in vivo
 CC or in vitro immunisation, particularly solid cancers and most especially
 CC hepatocarcinoma or adenocarcinoma of colon and kidney. They are also used
 CC to stimulate the immune system, and to increase, in culture, the
 CC production of associated-associated CTL, for reinfection, and/or to
 CC induce secretion of cytotoxic factors from CTL. Dendritic cells loaded
 CC with ICE are used to induce such CTL in cultures

XX Sequence 1680 BP; 345 A; 475 C; 488 G; 372 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,54e-106	Length:	1680
Score:	1230.00	Matches:	248
Percent Similarity:	63.84%	Conservative:	91
Best Local Similarity:	46.70%	Mismatches:	168
Query Match:	4	Indels:	24
		Gaps:	8

US-10-023-515-2 (1-581) x AAF25258 (1-1680)

```

QY 31 G1YProSerAlaG1YProGlnArgLeuThrArgLeuGlyTrp1LeGlnGlyLysGln 50
Db 76 GGCCAGAGACTCAGCCAGTCCATCCGAGCACACACGCGGACAGTGTGGAGTCTT 135
QY 51 ValThrValLeuGlySerProValProValAlaValPheLeuGlyValProPheAla 70
Db 136 GTCCATGTGAAGGGGCGCAATGCCGGGGTCCAAACCTTCCTGGAAATTCATTGCCAAG 195
QY 71 ProProLeuGlySerLeuThrPheThrAsnProGlnProAlaSerProThrAsnLeu 90
Db 196 CCACTCTAGTCCGCTGCGATTCACACCCCTGAGCCCTGAATCTTGAAGTGTGTG 255
QY 91 ArgGlnAlaThrSerTyProAsnLeuCyLeuGlnAsn-----SerGln 105
Db 256 AGGATGAGAACCAACCATCCGGCATGTGTCTACAGACCTCACCGCAGTGGAGTCAAG 315
QY 106 TrpLeuLeuLeuAspGlnHisMetLeuValHisTyProLysPheGlyValSerGln 125
Db 316 TTCTCT-----AGCCAGTTCAACATGACCTTCCTCCGACCTCATGCTGAG 363
QY 126 AspCyLeuTyLeuAsn1LeTyAlaProAlaHisAlaAspThrGlySerTyLeuPro 145
Db 364 GACTGCCCTGACCTCAGATCTACACGCGGCCCATAGCCATGAAGGCTCTAACTGCG 423
QY 146 ValLeuValThrPheProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGly 165
Db 424 GTGATGTGTGATCCACGATGCTGCGCTGTTTTCATGCGCTTCCTGTATGATGCT 483
QY 166 SerAlaLeuAla1TyArgLysAspValLeuValAlaGlnTyArgLeuGlyIle 185
Db 484 TCCATGCTGGCTGCTGGAGAACGTGGTGTGTCATCATCCGCTGAGTGC 543
QY 186 PheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheAspGln 205
Db 544 CTGGGCTTCTTACACACTGAGACAGACCAACCGCACTGGGCTTACCTGAGACAA 603
QY 206 ValAlaAlaLeuSerTrpValGlnLysAsn1LeGlnPheGlyGlyAspProSerSer 225
Db 604 GTGGCTTCACATCCCTGGGTCCACAGAAATATCCCACTTGGAGGACAACTGACCGT 663
QY 226 ValThrIlePheGlyGlyLeuSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPro 245
Db 664 GTACCAATTTTGGCGAGTCTGCGGGTGGCCAGAGTGTCTTGTGTGTGTCCTCC 723
QY 246 MetAlaLysGlyLeuPheHisTyAlaIleMetGlySerGlyValAlaIleIleProTy 265
Db 724 ATATCCCAAGGACTCTTCCACGAGCATATGAGAGTGGCGTGGCTCTCTGCTCCGCG 783

```

```

QY 266 LeuGlnAlaHisValArgTyGlnLysSerGlyAspLeu-----GlnValAlaAlaHisPhe 283
Db 784 CTAATTTGCC-----AGCTCAGCTATGATCATCTCCAGAGTGGGCCAACTGG 831
QY 284 CysGlyAsnAlaHisAspSerGlnAlaLeuLeuAlaGlyCyLeuAlaGlyThrLysProSer 303
Db 832 TCTGCTGTGACCAAGTTCAGTCTGAGGCTGTGGGTGCTGCGGGGCAAGATGA 891
QY 304 LysGlyLeuLeuThrLeuSerGlnLysThrGlySerPheThrArgValAlaAspGlyAla 323
Db 892 GAGAGAGATTTTGCATTAACAAGCCTTCAAGATATCCCGGAGTGTGATGGGCT 951
QY 324 PhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheValAlaIleProSer 343
Db 952 TTCTGCCCCAGCAACCCCAAGAGCTGTGCTGCTCCGACCTTTCAGCCCTGCTTAC 1011
QY 344 IleIleGlyValAlaAsnHisGlyGlyGlyPheLeuLeuPro-----MetLys 359
Db 1012 ATTGTGTGTTCACACAAATGAATTTGGCTGCTCATCCCAAGTCAATGAGATTTAT 1071
QY 360 GluAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuIleLeuGlnAsn 379
Db 1072 GATACCCAGAAAGAAATGACAGAGAGGCTCCAGAGCTGCTGTGCAAAATGTAAAG 1131
QY 380 IleLeuHisIleProProGlnTyLeuHisLeuValAlaAsnGlnTyPheHisValPheLys 399
Db 1132 CTGCTGATGTGCTCTTCAATTTGTGACCTGTGTGAGGAGAGTCAATTTGGGACAT 1191
QY 400 HisSerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 419
Db 1192 GGGATGCCCAAGACCTCCAAAGCCACTCCAGAGATGATGGCGACTCAGATTGTGTG 1251
QY 420 ValProAlaLeuIleThrAlaArgTyHisArgPheAlaGlyAlaProValTyPheTy 439
Db 1252 ATCCCTGACSTCCAAAGTACACATTTT---CAGTGTCCCGGGCCCTGTGTACTTAC 1308
QY 440 GluPheArgHisArgProGlnTyPheGluAspThrLysProAlaPheValIleValAsp 459
Db 1309 GAGTTCAGATCAACGCCAGCTGCTCAGAAACATCAGGCCACCGCAATGAAGGCGAG 1368
QY 460 HisAlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMet 479
Db 1369 CATGTGATGATGACTCTCTTTGTTTC---AGAAATTCTTTGGGGGCAACTATTA 1425
QY 480 PheGluGlyAlaThrGluGlnGlyLysLeuLeuSerArgLysMetLeuTyTrpAla 499
Db 1426 TTC-----ACTGAGGAAGAGAGCAGCTAAAGAGAAATGATGAACTGCGGCC 1476
QY 500 ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuThrProAlaTyra 519
Db 1477 AACTTTGCGAAGAAATGGAAACCCCAATGCGAGAGTCTCCACACTGCGCGCTGTTC 1536
QY 520 LeuThrGluGlnTyLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuGlyGln 539
Db 1537 CAGAGAGAGCAATACTGCAAGCTGAACCTACAGCTGCGGGGCGGCTGTGAAGGCC 1596
QY 540 ProArgValAspPheThrPheSerThrIlePro 550
Db 1597 CACAGGCTCAAGTTCTGAAAGAAAGCGCTGCC 1629

```

Search completed: June 15, 2005, 10:01:25
 Job time : 910 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 15, 2005, 12:02:07 ; Search time 953 Seconds

(without alignments)
3779.280 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079
1 MPOGUTSASQWCFPLIQP.....PLSLTFLSLQPFPPFCAP 581

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cg2_1/USPTO/6664091/runat_14062005_133643_10569/app_query.fasta-1.775
-DB=Published Applications NA -OPMT=fastcap -SUFFIX=p2n.tmpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -SPART=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pcrc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=6664091 @CGN_1_1480 @runat_14062005_133643_10569
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/1/pubpna/ECTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cg2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cg2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cg2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cg2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cg2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cg2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
19: /cg2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
20: /cg2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
21: /cg2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq:*
22: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cg2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	100.0	1746	13 US-10-023-515-3	Sequence 3, Appli
2	3079	100.0	1746	13 US-10-023-515-3	Sequence 3, Appli
3	3079	100.0	2158	13 US-10-023-515-1	Sequence 1, Appli
4	3079	100.0	2158	13 US-10-023-515-1	Sequence 1, Appli
5	3079	100.0	2158	13 US-10-023-515-1	Sequence 1, Appli
6	2915	94.7	1728	18 US-10-451-168-40	Sequence 23, Appli
7	2900	94.2	2232	18 US-10-433-256-23	Sequence 20, Appli
8	2877	93.4	2092	17 US-10-104-047-249	Sequence 249, App
9	2607.5	84.7	1746	18 US-10-451-168-41	Sequence 41, Appli
10	2606	84.6	1962	17 US-10-094-749-736	Sequence 76, App
11	2597.5	84.4	1746	18 US-10-114-270-195	Sequence 195, App
12	2588.5	84.1	1857	18 US-10-381-698-20	Sequence 20, Appli
13	2015	65.4	2145	20 US-10-233-933A-3	Sequence 3, Appli
14	1985	64.5	1629	20 US-10-233-933A-1	Sequence 1, Appli
15	1780	57.8	1071	18 US-10-451-168-42	Sequence 42, Appli
16	1230	39.9	2191	9 US-09-980-107-3854	Sequence 3854, Ap
17	1230	39.9	2191	9 US-09-954-531-1038	Sequence 1038, Ap
18	1230	39.9	2191	10 US-09-873-367C-828	Sequence 828, App
19	1230	39.9	2191	10 US-09-873-367C-829	Sequence 829, App
20	1230	39.9	2191	20 US-10-858-271-27	Sequence 27, Appli
21	1230	39.9	2191	21 US-10-843-641A-828	Sequence 828, App
22	1230	39.9	2191	21 US-10-843-641A-829	Sequence 829, App
23	1230	39.9	2191	21 US-10-843-641A-2105	Sequence 2105, Ap
24	1230	39.9	2484	9 US-09-925-298-271	Sequence 271, App
25	1230	39.9	2484	14 US-10-102-806-271	Sequence 271, App
26	1215.5	39.5	1902	17 US-09-917-800A-480	Sequence 480, App
27	1215.5	39.5	1902	17 US-10-368-934-40	Sequence 40, Appli
28	1215.5	39.5	1902	18 US-10-152-319A-531	Sequence 531, Appli
29	1181	38.4	2087	17 US-09-895-860-1	Sequence 1, Appli
30	1181	38.4	2087	17 US-10-377-072-1	Sequence 1, Appli
31	1179.5	38.3	2092	21 US-10-956-157-4153	Sequence 4153, Ap
32	1179.5	38.3	3824	13 US-09-931-836-22	Sequence 22, Appli
33	1179.5	38.3	3824	13 US-10-036-442-22	Sequence 22, Appli
34	1179.5	38.3	3824	13 US-10-036-041-22	Sequence 22, Appli
35	1179.5	38.3	3824	13 US-10-028-072-541	Sequence 541, App
36	1179.5	38.3	3824	14 US-10-035-855-22	Sequence 22, Appli
37	1179.5	38.3	3824	14 US-10-140-808-541	Sequence 541, App
38	1179.5	38.3	3824	14 US-10-121-049-541	Sequence 541, App
39	1179.5	38.3	3824	14 US-10-123-904-541	Sequence 541, App
40	1179.5	38.3	3824	14 US-10-140-470-541	Sequence 541, App
41	1179.5	38.3	3824	14 US-10-175-746-541	Sequence 541, App
42	1179.5	38.3	3824	14 US-10-176-918-541	Sequence 541, App
43	1179.5	38.3	3824	14 US-10-176-921-541	Sequence 541, App
44	1179.5	38.3	3824	14 US-10-176-921-541	Sequence 541, App
45	1179.5	38.3	3824	14 US-10-227-884-209	Sequence 209, App

ALIGNMENTS

RESULT 1
US-10-023-515-3
; Sequence 3, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy A. J.
; APPLICANT: Sltos-Santiago, Inmaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYESTERASE
; FILE REFERENCE: 10448-122601
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28

Pred. No. is the number of results predicted by chance to have a


```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-515-3

```

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	3079.00	1746	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-10-023-515-2 (1-581) x US-10-023-515-3 (1-1746)

```

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleuGlnPro 20
Db 1 ATGCCACAGGAGCTTACTTTCATCTGCTTCACATGAGTGTCTTTCTGATTCTCCAGCCC 60
QY 21 LeuLeuGlyVHIAArgGlnTrpGlyVyrGlyProSerAlaGluGlyProGlnArgan 40
Db 61 CTGTGGGACACAGACAGTGGGGAAAACCTGGCCCTTCTGCTGAGGGCCACAGAGAAC 120
QY 41 ThrArgLeuGlyTrpIleGlnGlyVyrGlnValThrValLeuGlySerProValProVal 60
Db 121 ACCAGGCTGGAGTGGATTTCAGGGCAAGACATCTGTCTGGGAAGCCCTGTGCTGTG 180
QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
Db 181 AACGTGTTCTGTGGAGTCCCTTTGTCTGCCCGCTGGAGTCCCTGGATTTCAGAAC 240
QY 81 ProGlnProAlaSerProTrpPhePheLeuArgGlnValIleSerTrpProAlaLeuGly 100
Db 241 CGGACCTGTCATCGCCCTGGAGATTAACCTTGCAGAAAGCCACTTCACTTAAATTGTGC 300
QY 101 LeuGlnAsnSerGlyTrpLeuLeuLeuArgGlnHISMetLeuValValHISTrpProLys 120
Db 301 CTCGAGAACTCAGAGTGGGTGCTCTTAATCAACATGCTCAAGTGCATTACCCGAAA 360
QY 121 PheGlyValSerGlyAspCysLeuTrpLeuAsnIleTyraIleProAlaHISAlaAspThr 140
Db 361 TTCGAGAGTGCAGAACTGCTCTTAACCTGAACATCTATGCGCCCTGCACGCGCATCA 420
QY 141 GlySerIleLeuProValIleValIleTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
Db 421 GGCTCCAGGCTCCCGGCTTGTGTGTGTTCCAGAGAGGTGGCTTCMAAGCTGGCTCAGCC 480
QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTyrglyAspValIleValIleValIleGln 180
Db 481 TCCATCTTGAATGGGTCCGCCCTGGCTGCTATGAGGAGCGTGTGTGTGTGTGTGTGT 540
QY 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHISAlaProGlyAsnTrp 200
Db 541 TACGGCTGAGAAATATTTGTTCTTCACACATGAGGATCAGATGCTCCGGGGAACCTGG 600
QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGlnPhePheGly 220
Db 601 GCCTTCAGAGACAGAGTGGCTGTCTGTCTGGGTCCAGAAAGAACATGCAATTTCTGGGT 660
QY 221 GlyAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSer 240
Db 661 GGGGAGCCCAAGCTCTGTACATCTTTGGCGAGTCCGGGGAGCCATATGATGTTTCAAGT 720
QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHISLysAlaAlaIleMetGluSerGlyVal 260
Db 721 CTTATACTGTCTCCCAAGGCCAAAGGCTTATTCACAAAGCACAAGAGAGAGGGGTG 780
QY 261 AlaIleIleProTrpLeuGlnValAlaHISAspTrpGluLysSerGluAspLeuGlnValIle 280
Db 781 GCCATCATCCCTTACTCTGAGGAGCCCATGATTATGAAAGAGTAGAGACTGTGCAAGTGTG 840

```

```

QY 281 AlaHISpHeCySGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
Db 841 GCACATTTCTGTGTGTAAACAATGCGTCAAGCTCTGAGGCTGTGTGAGGTGCTGAGACA 900
QY 301 LysProSerIleVSGluLeuLeuThrIleuSerGlnLysThrLysSerPheThrArgValVal 320
Db 901 AAACCTTCMAAGAGACTGCTGACCTCAGCCAGAAACMAAGTCTTTTCACTGAGTGGTT 960
QY 321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysValAlaPheLysVal 340
Db 961 GATGCTGCTTCTTCTTCTTAAGAGCTTCAATCTATTGTCTTCAGAAAGCATTTAAAGCA 1020
QY 341 IleProSerIleIleGlyValAlaAsnAsnISGluCySGlyPheLeuLeuProMetLysGlu 360
Db 1021 ATTCTTCATCATCAGAGTCAATTAACACAGAGTGTGCTTCTGTGCTGTGCTATGAAGAG 1080
QY 361 AlaProGluIleuSerGlySerAsnLysSerLeuAlaLeuHISLeuIleGlnAsnIle 380
Db 1081 GCTCTGAGATCTCAAGTGGCTCCAGAACAGTCCCTTGCCCATCTGATTAACAAACATC 1140
QY 381 LeuHISIleProProGlnTrpLeuHISLeuValAlaAsnGluTrpPheHISAspLysHIS 400
Db 1141 CTGACATCCCGGCTCAGTATTTCACCTTGTGGCTTAATGAATCTTCCATGACAAAGAC 1200
QY 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
Db 1201 TCCCTGACTGAATCCAGACAGTCTTGGACTTGTGGAGATGTGTCTTGTGTGTG 1260
QY 421 ProAlaLeuIleThrAlaArgTrpHISArgAspAlaGlyAlaProValTrpPheTrpGlu 440
Db 1261 CTTCACATGATCAGATCTCATATACAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
QY 441 PheArgHISArgProGlnCyPheGluAspThrLysProAlaPheValLysAlaAspHIS 460
Db 1321 TTTCGGACCGGCTCAGTCTTGAAGACAGGAGCCGCTTTTGTCAAAAGCGAGAC 1380
QY 461 AlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValIlePhe 480
Db 1381 GCTATAGATGTCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
QY 481 GlnGlyValIleThrGlnGlnGlyLysLeuLeuSerArgLysMetLysTrpTrpAlaThr 500
Db 1441 GAAGAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyraLeu 520
Db 1501 TTGTCTGAAACCGGGAATCTTAATGGGAAAGACCTGTCTGTGTGTGTGTGTGTGTGT 1560
QY 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
Db 1561 ACTGAGCAATCTTCCAGCTGAGCTTGAACATGAGCTTGGAGAGAGACTCAAAAGAACG 1620
QY 541 ArgValAspPheThrPheSerThrIleProLeuIleLeuSerAlaSerAspMetLysHIS 560
Db 1621 CGGGTGAATTTTGAACAGACACATCCCGTGAATCTGTGTGTGTGTGTGTGTGTGTGT 1680
QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCyVala 580
Db 1681 AGTCTCTTCTTCTTAACTTTCCTCTCTCTCTCCACACTTCTTCTTCTTCTTGTGT 1740
QY 581 Pro 581
Db 1741 CCT 1743

```

RESULT 2

```

US-10-674-636-3
; Sequence 3, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; APPLICANT: Siles-Santlago, Imaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE

```

```

1  TITLE OF INVENTION:  FAMILY MEMBER AND USES THEREOF
2  FILE REFERENCE:  10448-122001
3  CURRENT APPLICATION NUMBER:  US/10/0,674,636
4  CURRENT FILING DATE:  2003-09-29
5  PRIOR APPLICATION NUMBER:  US/10/023,515
6  PRIOR FILING DATE:  2001-12-18
7  PRIOR APPLICATION NUMBER:  60/256,369
8  PRIOR FILING DATE:  2000-12-18
9  PRIOR APPLICATION NUMBER:  60/219,508
10 PRIOR FILING DATE:  2001-03-28
11 NUMBER OF SEQ ID NOS:  6
12 SOFTWARE:  FastSeq for Windows Version 4.0
13 SEQ ID NO 3
14 LENGTH:  1746
15 TYPE:  DNA
16 ORGANISM:  Homo sapiens
17 US-10-674-636-3

```

Alignment Scores:		
Pred. No.:	0	Length: 1746
Score:	3079.00	Matches: 581
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	18	Gaps: 0

Qy	1	MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuLeuGlnPro	20
Db	1	ATGCCACAGGAGCATTAATCTCAATCTGCTTCAACAATGGAGCTTTTCTGATCTTCACAGCC	60
Qy	21	LeuLeuGlnYhiSarGlnTrpGlyLeuYsrThrGlyProSerAlaGlnGlyProGlnAtgAsn	40
Db	61	CTGTGTGGACACAGACAGTGGGGAAAACTGGGCTTTCGTGTGAAGGGCCACAGAGAAC	120
Qy	41	ThrArgLeuGlnYTrpLleGlnGlyYsrGlnValThrValLeuGlySerProValProVal	60
Db	121	ACCAAGCTGGAGATGATTCAGGACCAAGATCATGTGTGTGGAAAGCCCTGTGCTGTG	180
Qy	61	AsnValPheLeuGlnValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn	80
Db	181	AACGTGTTCTCTGGAGTCCCTTTGTGTGCTCCCGCTGGAGTCCGTGGCATTTTAGAAC	240
Qy	81	ProGlnProAlaSerProTrpPheAsnLeuArgGlnAlaThrSerTrpProAsnLeuCys	100
Db	241	CCGCAAGCTGCATCCCGCTGGAGATTACTTGGCGAAGGCCACCTCTTAACCTAATTTGTGC	300
Qy	101	LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuYsrValHisTrpProIys	120
Db	301	CTCCGAATCTAGATGGTGGCTGCTCTTAGATCAACAATGCTCAAGTGCAATTACCGGAA	360
Qy	121	PheGlyValSerGlyAspCysLeuTrpLeuAsnLietrAlaProAlaHisAlaAspThr	140
Db	361	TTCCGAGTGTCAAGAGACTGCCTCACTGAACATCTAATGCCCTGCCACGCCGATACA	420
Qy	141	GlySerTrpLeuProValLeuValTrpPheProGlyGlyValaPheYsrThrGlySerAla	160
Db	421	GGCTTCAAGCTCCCGCTCTGTGTGTGGTTCACAGAGGTGCTTCAAGATGTGCTACGCC	480
Qy	161	SerLlePheAspGlySerAlaLeuAlaAlaTrpGlyAspValLeuValValAlaGln	180
Db	481	TCCATCTTTGATGGGTCCGCCCTGGGTGCTATGAGAGACGTGTGTTGTGTGCTGCCAG	540
Qy	181	TyrArgLeuGlyLlePheGlyPhePheThrTrpTrpAspGlnHisAlaProGlyAsnTrp	200
Db	541	TACCCGCTAAGAAATATTTGGTTCTTCAACACATGGAGTACAGATCTCTCGGGGAACTGG	600
Qy	201	AlaPheYsrAspGlyValAlaAlaLeuSerTrpValGlnYsrAsnLleGlnPhePheGly	220
Db	601	GCCTTCAGAGACAGGTGCTGCTGTCTCTGGGTTCAGAGAACATCGAGTTCTTGGGT	660
Qy	221	GlyAspProSerSerValThrLlePheGlyGlnSerAlaGlyAlaLleSerValSerSer	240

[illegible]

```
Db          1741 CCT 1743
RESULT 3
US-10-023-515-1
; Sequence 1, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; PRIORITY REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023, 515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-023-515-1
Alignment Scores:
Pred. No.: 0
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Length: 2158
Matches: 581
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-023-515-2 (1-581) x US-10-023-515-1 (1-2158)
QY      1 MetProGlnGlyLeuThiSerSerAlaSerGlnTrpCysPhePheLeuIleuGlnPro 20
Db      96 ATGCCACAGAGGACCTTCACTTCCTGCTTCAACATGAGTGGCTTTCTCTGATTCCTCAGGCC 155
QY      21 LeuLeuGlyYHIArGlnTrpGlyLeuThiSerProSerAlaGlnGlyProGlnAraGln 40
Db      156 CTGTTGGGACACAGACAGTGGGAAACTGGGCTTCTGCTGAAGGGCCACAGAGGAAAC 215
QY      41 ThrAlaGlnGlyTrpIleGlnGlyLeuGlnValThrValLeuGlySerProValProVal 60
Db      216 ACCAGAGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAAAGCCCTGTGCTGTG 275
QY      61 AenValPheLeuGlyValProPheAlaAlaProProlLeuGlySerLeuAraGlnPheThrAsn 80
Db      276 AACGTTCTCTCGAGGATCCCTTGTGCTCTCCCGCTGGGATCCCTGCGATTATTCAGAAC 335
QY      81 ProGlnProAlaSerProTrpAraPheLeuAraGlnAlaThrSerTrpProAenLeuCys 100
Db      336 CCGGAGCTGATCGCCCTGGGATTAACCTTGGGAAAGCACTCTTAACCTTAATTTGTGC 395
QY      101 LeuGlnAenSerGlyTrpPheLeuLeuAraGlnIlePheLeuValIleGlyProGly 120
Db      396 CTCGAAACCTAGAGTGGCTCTTGTAGATCAACAGATGCTCAAGGTGCAATTACCGGAA 455
QY      121 PheGlyValSerGlyAraPheCysLeuTrpLeuAenIleTrpAlaProAlaAlaIleAraPheThr 140
Db      456 TTGCGAGGTGTAGAAAGCTGCTTAACCTTAACCTTAAGGCTGCGCAAGCGCGATACA 515
QY      141 GlySerLeuLeuProValLeuValTrpPheProGlyGlyValAraPheLeuThiSerAla 160
Db      516 GGGCCCAAGCTCCCGCTTGTGTGTGCTCCAGAGAGTGCCTTCAAGACTGAGCTGAGCC 575
QY      161 SerIlePheAraPheGlySerAlaLeuAlaAlaTrpGlyAraPheValLeuValValGln 180
Db      576 TCCATCTTTGATGGGTCCGCCCTGGCTGCTATGAGAGAGTGTGTTGTGTGCTCCAG 635
QY      181 TyrArgLeuGlyIlePheGlyPhePheThrTrpAraGlnIleAlaProGlyAraTrp 200
Db      636 TACCGCTAGGAATTTTGGTTTCTTCCACATGAGATCAGCAAGCTCCGGGAAACGCG 695
QY      201 AlaPheLeuAraPheGlnValAlaAlaLeuSerTrpValGlnLeuAenIleGlyPhePheGly 220
Db      696 GCTTTCAGAGACCAAGTGGCTGCTGTCTGGGGTCCAGAAAGAACATCGAATTTCTTGGT 755
QY      221 GlyAraPheProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
Db      756 GGGAGCCCAAGCTGTGATCAATCTTTGGCAGATCCCGGGAGCCATTAAGTGTTCAGT 815
QY      241 LeuIleLeuSerProMetAlaLeuGlyLeuPheHisValAlaIleMetGlnSerGlyVal 260
Db      816 CTTATACGTCTCCCATATGCCAAAGGCTTAATCCAAAGCCATCATGAGAGAGTGGGTT 875
QY      261 AlaIleIleProTrpLeuGlnAlaAlaPheTrpGlnLeuSerGlyAraPheLeuGlnVal 280
Db      876 GCCATCATCCCTTACCTGGAGGCCCATGATTAAGAAAGAGTGAAGAGCTGCAAGTGTGT 935
QY      281 AlaHisPheCysGlyAraAenAlaSerAraPheSerGlnAlaLeuLeuAraGlySerLeuAraGln 300
Db      936 GCACATTTCTGTGTGTTAACAATGCTGTCAAGCTCTGAGGCCCTGCTGAGGTCTTGAGACA 995
QY      301 LeuProSerLeuGlyLeuLeuLeuThiLeuSerGlnLeuThiLeuSerPheThrAraGlyVal 320
Db      996 AAACCCCTCAAGAGAGCTGAGACCTCAGCCACAGAAACAAAGCTTTCACCTGAGAGTGT 1055
QY      321 AarGlyAlaPhePheProAenGlnProLeuAraPheLeuSerGlnValAlaPheLeuVal 340
Db      1056 GATGCTCTTCTTCTTCTTAATAGCCCTTACATCTTATGTCTCAGAAAGCATTTAAAGCA 1115
QY      341 IleProSerIleIleGlyValAaAenAlaGlnCysGlyPheLeuLeuProMetLeuGln 360
Db      1116 ATTCTTCCATCATCGGAGTCATTAACCAAGATGTGTGCTTCTGCTGCTATGAAGAG 1175
QY      361 AlaProGlnIleLeuSerGlySerAenLeuSerLeuAlaLeuIleGlnAenIle 380
Db      1176 GCTCTGAGATCTCTCAGTGGCTTCCAAAGTCCCTTGGCCCTCATCTGATACAAACATC 1235
QY      381 LeuHisIleProProGlnTrpLeuHisLeuValAlaAenGlyTrpPheHisAraPheHis 400
Db      1236 CTCGACATCCCGCTCAGTATTTGACCTTGTGGCTAATGATCTTCAATACAAAGAC 1295
QY      401 SerLeuThiGlnIleAraPheSerLeuAraPheLeuGlyAraPheValPhePheVal 420
Db      1296 TCCCTAAGTAATCCGAGACAGTCTTGTGACTTGTGGAGATGTCTTGTGTGTC 1355
QY      421 ProAlaLeuIleThrAlaArgTrpHisAraPheAlaGlyAlaProValTrpPheTrpGln 440
Db      1356 CTGCACTGATACAGCTCGATATCAACAGATGTGTGGGACCTGTCTACTCTTATAG 1415
QY      441 PheArgHisAraProGlnCysPheGlyAraPheTrpLeuProAlaPheValAlaAraPheHis 460
Db      1416 TTTCCGCAACCGGCTCAGTCTTGAACAACAGAGCCGGCTTTTGTCAAAAGCCGACAC 1475
QY      461 AlaAraGlyValAraPheValPheGlyGlyAlaPheLeuLeuGlyAraPheIleValIlePhe 480
Db      1476 GCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1535
QY      481 GlnGlyAlaIleGlnGlnGlnGlnLeuLeuSerArgLeuSerMetLeuSerTrpAlaThr 500
Db      1536 GAAAGAGCCACGAGAGAGAGAGATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1595
QY      501 PheAlaArgTrpGlyAraProAenGlyAraAraPheLeuSerLeuTrpProAlaTrpAraPhe 520
Db      1596 TTTGCTCAACCGGAGAACTTAATGGGAAACGACTGTGTGTGTGTGTGTGTGTGTGTGTGT 1655
QY      521 ThrGlnGlnTrpLeuGlnLeuAraPheAraPheSerLeuGlyGlnAraGlyLeuValPro 540
Db      1656 ACTGAGAGTACCTCTCAGAGTGTGACTTGAACATGAGCTCCGAGCAGAGACTCAAGAGAGCCG 1715
```

QY 541 ArgValaepheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
DB 1716 CCGGGGATTTTGGACGACGACACCTCCGATCTCTGTGCTCCGACATGCTCCAC 1775
QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
DB 1776 AGTCTCTTTCTTCTTAACCTTCTCTCTCTCTCCGACCTTTCTTTTCTTTGTGCT 1835
QY 581 Pro 581
DB 1836 CCT 1838
RESULT 4
US-10-674-636-1
; Sequence 1, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; FILE REFERENCE: 10448-12201
; CURRENT APPLICATION NUMBER: US/10/674, 636
; PRIOR FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023, 515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279, 508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-674-636-1
Alignment Scores:
Pred. No.: 0 Length: 2158
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-10-023-515-2 (1-581) x US-10-674-636-1 (1-2158)
QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
DB 96 ATGCCACAGGAGCATTAATCTCATCTGCTTCAACAATGCTGCTTTTCTGATTTCTCCAGCCC 155
QY 21 LeuLeuGlnGlyHisArgGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnArgan 40
DB 156 CTGTTGGGACACAGACAGTGGGAAAAAATGGGCTTCTGCTGAAGGGCCACAGAGGAAC 215
QY 41 ThrArgLeuGlyTrpIleGlnGlyLeuGlnValThrValLeuGlySerProValProVal 60
DB 216 ACCAGGCTGGAGTGAATTCAGGCGACAGTCACTGCTGGAGAGCCCTGTGCTGTG 275
QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
DB 276 AACGTGTTCTCGAGAGTCCCTTTGCTGCTCCCGCTGGATCCCTGGATTACAGAC 335
QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnValThrSerTrpProAsnLeuCys 100
DB 336 CCGCAGCGCTGATCGCCCTGGGATTAATTCGCGAAGAGCCACCTCTTAATTTGTGCT 395
QY 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuValValHisTrpProLys 120

DB 396 CTCGAGAACTAGAGTGGCTGCTTAATGATCAACATGCTCAAGTGTGATTAACCGAAA 455
QY 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
DB 456 TTCCGAGAGTCAAGAGACTGCTCTTAATCTGAACATCTTATGCCCCCGCCACCGCATTA 515
QY 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
DB 516 GGCTCCAGACTCCCGCTTGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGCTCAGGC 575
QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlyAspValLeuValValGln 180
DB 576 TCATCTTTGAGTGGATCCGCCCTGGCTGCTGATGAGAGTGTCTGTGTGCTGCTGAC 635
QY 181 TrpArgLeuGlyTrpPheGlyPhePheThrTrpAspGlnHisAlaProGlyLeuTrp 200
DB 636 TACCGGCTAAGAAATTTGTGTTCTTCAACATATGGAGTCAAGCAATGCTCCGGGAACTGG 695
QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGlyPhePheGly 220
DB 696 GCCTTCAAGACCAAGTGGCTGCTGTCTGTGGTCCAGAGAACATCGAGTTCTTGGT 755
QY 221 GlyAspProSerSerValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSer 240
DB 756 GGGGACCCAGCTCTGTGACCATCTTTGGCGAGTCCGGGGAGCCATAAGTGTCTACT 815
QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisLysValIleMetGlySerGlyVal 260
DB 816 CTTAATAGTCTCCCATGCGCCAAAGCTTATCCAAAGCATCATGAGAGTGGGGTG 875
QY 261 AlaIleIleProTrpLeuGlnValHisAspTrpGlyLysSerGlnAspLeuGlnValVal 280
DB 876 GCCATCATCTCTTACCTGAGGCGCATGATTAAGAAAGAGAGAGCACTGAGGTGTT 935
QY 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
DB 936 GCACATTTCTGTGATGATGATGCGTCAGACTGTGAGGCGCTGAGGGTCTGAGGACA 995
QY 301 LysProSerLysGlnLeuLeuThrLeuSerGlnLysTrpLysSerPheThrArgValVal 320
DB 996 AAACCTCCAGAGACTGCTGACCTCCAGCCAGAAACAAAGTCTTCACTGAGTGGTT 1055
QY 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuSerGlnLysAlaPheLysAla 340
DB 1056 GATGAGTCTTCTTCTTAATGAGCTCTAGATCTTAATGTCTCAAGAGCAATTAAAGCA 1115
QY 341 IleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLysGln 360
DB 1116 ATTCCTTCATCATGAGATCAATAACACAGATGCTGCTCTGCTGCTATGAAGAG 1175
QY 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIle 380
DB 1176 GCTCTGAGATCTCAGTGGCTCCAAAGTCCCTTGCCTCATCTGATCAAAAACATC 1235
QY 381 LeuHisIleProProGlnTrpTrpLeuHisLeuValAlaAsnGlnTrpPheHisAspLysHis 400
DB 1236 CTGACATCTCCCGCTCAGATTTTGCACCTTGAGCTAATGAATCTTCATTAACAAGCAC 1295
QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
DB 1296 TCCCTGATCGAAATCCGAGACAGTCTTCTGACCTGCTGGAGAGTGTCTTTGTGGTTC 1355
QY 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGlu 440
DB 1356 CTGCACTGATCAAGCTCGATATCAAGAGATGCTGTGGTGCACCTGTACTTATATGAG 1415
QY 441 PheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspHis 460
DB 1416 TTTCGAGACCGGCTCAGTGTGTTGAAGACAGAAAGCGGCTTTGTCAAAAGCCACAC 1475
QY 461 AlaAspGlyValAlaArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480

```

FEATURE:
; NAME/KEY: CDS
; LOCATION: (96) ... (1841)
US-10-757-262-45

Alignment Scores:
Pred. No.: 0
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 19

Length: 2158
Matches: 561
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-023-515-2 (1-581) x US-10-757-262-45 (1-2158)

```

QY	MeProGI ₁ NGI ₁ YeuThrsSer ₁ AlaSer ₁ INT ₁ PCy ₁ PhePheLeu ₁ IleuGI ₁ Pro ₂₀
Db	ATGCACAGGAGCTTACTTCTCATCTGCTTCACACATGATGCTTTTCTTATATCTCCAGGCC ₁₅₅
QY	21 LeuLeuGI ₁ YH ₁ ArgGI ₁ INT ₁ PCy ₁ YthrThrgI ₁ ProSer ₁ agiGI ₁ YProGI ₁ ArgAn ₄₀
Db	156 CTGTTCGGGACACAGCACAGTGGGGGAAACCTGGGCTTCTGCTGAAGGGCCACAGAGGAC ₂₁₅
QY	41 ThrArgLeuGI ₁ TrpIleGI ₁ NGI ₁ YysGI ₁ Val ₁ ThrVal ₁ IleuGI ₁ YserPro ₁ Val ₆₀
Db	216 ACCAGGCTGGATGGATTTCAGGGCAAGCAAGTCACTGCTGGGAAGACCTGTGCTGGTG ₂₇₅
QY	61 AsnVal ₁ PheLeuGI ₁ Val ₁ ProPhe ₁ Ala ₁ AlaPro ₁ LeuGI ₁ YserLeuArgPheThrAsn ₈₀
Db	276 AACGTGTCCTGGAGTCCCTTTTGCTCTCCCGCGGTGGATCCCTGGATTTCAGAAC ₃₃₅
QY	81 ProGI ₁ ProAlaSer ₁ ProTrpAspAsnLeuArgI ₁ Val ₁ ThrSer ₁ Trp ₁ ProAsnLeu ₁ Cys ₁₀₀
Db	336 CCGCAGCTGCATCCGCCCTTGGATACCTTGGAGAAGCACCTCTACCCCTAA ₃₉₅ TTTGTC
QY	101 LeuGI ₁ AsnSer ₁ GI ₁ TrpLeuLeuLeuAspGI ₁ NI ₁ smcLeu ₁ YsVal ₁ His ₁ Trp ₁ Pro ₁ Ys ₁₂₀
Db	396 CTCCTCAAACTCAGATGGCTGCTCTTAGATCAACACATGCTCAAGTGCACTTACCGAAA ₄₅₅
QY	121 PheGI ₁ Val ₁ SerGI ₁ AspCysLeu ₁ TrpLeuAsn ₁ Ile ₁ TrpAlaProAlaHis ₁ AlaAspThr ₁₄₀
Db	456 TTCGAGGTGTCAGAAAGACTGCTCTTACCTGAACATCATGCGCTGCCACGCCGATAC ₅₁₅
QY	141 GlySer ₁ YsLeuPro ₁ Val ₁ LeuVal ₁ TrpPhePro ₁ GI ₁ Val ₁ AlaPhe ₁ YthrThrgI ₁ YserAla ₁₆₀
Db	516 GGCCTCAAGACTCCCGCTGCTGTGTGGTTCACAGAGAGCTCCTTCAAGACTGCTCAGCC ₅₇₅
QY	161 SerIlePheAspGI ₁ YserAlaLeuAla ₁ Ala ₁ TrpGI ₁ AspVal ₁ LeuVal ₁ Val ₁ Gln ₁₈₀
Db	576 TCCATCTTATATGGAGTCCCGCTGGCTGACCTTAAGAGACCTGCTGGTTGGTGGTCCAG ₆₃₅
QY	181 TyrArgLeuGI ₁ IlePheGI ₁ PhePheThr ₁ Thr ₁ TrpAspGI ₁ NI ₁ His ₁ Ala ₁ ProGI ₁ YAsn ₁ Trp ₂₀₀
Db	636 TACGGGCTAAGGAATATTGGTTTCTTACACATGGGATTCAGCATGCTCCGGGAACTGG ₆₉₅
QY	201 AlaPheYsAspGI ₁ Val ₁ Ala ₁ AlaLeuSer ₁ TrpVal ₁ Gln ₁ YsAsn ₁ IleGI ₁ upPhePheGI ₁ ₂₂₀
Db	696 GCTTTCAGAGCACAGGTGGCTGCTGTCTCTGGGGTCCAAAGAACATCGAGTTCTTCGGGT ₇₅₅
QY	221 GlyAspProSerSerVal ₁ ThrIlePheGI ₁ YsSer ₁ GI ₁ Val ₁ Ala ₁ IleSer ₁ Val ₁ SerSer ₂₄₀
Db	756 GGGGACCCAGCTGTGTGACATCTTTGGCGAGTCCGGCGAGCCCAATAGTATTTCTTCAGT ₈₁₅
QY	241 LeuIleLeuSerProMetAla ₁ Arg ₁ LeuPheHis ₁ YsAla ₁ Ala ₁ MetGI ₁ YsSerGI ₁ Val ₁ ₂₆₀
Db	816 CTTATACCTGTCTCCCATGCGCAAAAGCTTATTTCCAAAGGCCATCATGAGATGGAGTGGGGTG ₈₇₅
QY	261 AlaIleIlePro ₁ TrpLeuGI ₁ Val ₁ His ₁ Asp ₁ TrpGI ₁ YsSerGI ₁ AspLeuGI ₁ Val ₁ ₂₈₀
Db	876 GCCATCATCCCTTACCTCGAGGCCCAATGATTTATGAGAAAGTGAAGACCTGCAAGGTGGTT ₉₃₅
QY	281 AlaHisPheCysGI ₁ YsAsnAsn ₁ AserAspSerGI ₁ Val ₁ AlaLeuArgCysLeuArgThr ₃₀₀

```

Db 936 GCACATTTCTGTGTAACATGCTCAGACTCTGAGCCCTGAGTGCCTGAGGACA 995
Qy 301 LysProSerLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
Db 996 AAACCTTCACAGAGAGCTGTGACCTTCAGCCAGAAACAAAGTTCTTCACTGAGTGT 1055
Qy 321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
Db 1056 GATGTGCTTCTTCTTCTTAATGAGCCTCAGATCTATTGTCTCAGAAAGCATTTAAAGCA 1115
Qy 341 IleProSerIleIleGlyValAsnAsnHisGluCysGlyPheLeuLeuProMetLysGlu 360
Db 1116 ATTCCTTCATCATCATGAGTCAATTAACCAAGATGTGCTCTCTGCTCATGAAAGAG 1175
Qy 361 AlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisIleuIleGlnAsnIle 380
Db 1176 GCTCTGAGATCTCAGTGGCTCCCAACAGTCCCTTGGCTTCATGATTAACAAACATC 1235
Qy 381 LeuHisIleProProGlnIleThrLeuHisIleuValAlaAsnGluThrPheHisAspLysHis 400
Db 1236 CTGCACATCCCGCTCAGATTTGCACTTGTGCTATGATTAATTCTTCATGACAAAGCAC 1295
Qy 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
Db 1296 TCCCTGACTGAATCCGAGACAGTCTTCTGACTTGTGAGATGTGTCTTGTGTGTC 1355
Qy 421 ProAlaLeuIleThrAlaArgThrHisArgAspAlaGlyAlaProValThrPheThrGlu 440
Db 1356 CCTGACATGATCAGCTCGATATCAGAGATGCTGTGACCCGTCTACTCTATGAG 1415
Qy 441 PheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspHis 460
Db 1416 TTTCGGCACCCGGCTCAGTCTTTTAAGACACGAAAGCCGGCTTTGTCTCAAGCCCAACAC 1475
Qy 461 AlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
Db 1476 GCTGATGAAGTCCGCTTGTGTTCGATGGTGCCTTCGAAGGGGAGCAATTGTATGTTTC 1535
Qy 481 GluGlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysThrTrpAlaThr 500
Db 1536 GAAGGAGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATAATATCAGGGCTACCC 1595
Qy 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaThrAsnLeu 520
Db 1596 TTTCGTCACACCGGAATCTTAATGGAACGACTGTCTCTGTGCGCACCTTAATATCTG 1655
Qy 521 ThrGluGlnIleThrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
Db 1656 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCTTCGAGCAGAGACTCAAAAGACCG 1715
Qy 541 ArgValAspPheThrThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
Db 1716 CGGGGTGATTTTGTGACACGACCATCCCTCGATCTGTCTGCTCCGACATGCTCCAC 1775
Qy 561 SerProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePhePheCysAla 580
Db 1776 AGTCTCTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
Qy 581 Pro 581
Db 1836 CCT 1838

```

```

/ CURRENT FILING DATE: 2003-11-12
/ PRIOR APPLICATION NUMBER: PCT/US01/49232
/ PRIOR FILING DATE: 2000-12-17
/ PRIOR APPLICATION NUMBER: 60/256,710
/ PRIOR FILING DATE: 2000-12-19
/ PRIOR APPLICATION NUMBER: 60/257,048
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 60/260,482
/ PRIOR FILING DATE: 2001-01-09
/ PRIOR APPLICATION NUMBER: 60/264,922
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/266,797
/ PRIOR FILING DATE: 2001-02-06
/ PRIOR APPLICATION NUMBER: 60/276,988
/ PRIOR FILING DATE: 2001-03-19
/ PRIOR APPLICATION NUMBER: 60/281,535
/ PRIOR FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: 60/289,622
/ PRIOR FILING DATE: 2002-06-28
/ NUMBER OF SEQ ID NOS: 110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 40
/ LENGTH: 1728
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-451-168-40

Alignment Scores:
Pred. No.: 0 Length: 1728
Score: 2915.00 Matches: 553
Percent Similarity: 97.54% Conservative: 3
Best Local Similarity: 97.02% Mismatches: 6
Query Match: 94.67% Indels: 8
D: 18 Gaps: 1

US-10-023-515-2 (1-581) x US-10-451-168-40 (1-1728)
12 TrpCysPhePheLeuIleLeuGlnProLeuLeuGlyHisArgGlnTrpGlyLysThrGly 31
Db 40 TGGGCTATCTGGGTCTTTCAGACCCCTC-----ACCAAGGG 75
Qy 32 ProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLysGlnVal 51
Db 76 CTTTCTGTGAAGGGCCACAGAGAAACCAAGCTGGATGATTCAGGGCAAGCAAGTTC 135
Qy 52 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAlaPro 71
Db 136 ACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTCTCTCGAGAGTCCCTTGTGCTGCC 195
Qy 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArg 91
Db 196 CCGCTGGATCCCTCGCATTTACGAACCCGAGCCCTGATGCTCCCTGGATTAACCTTGGCA 255
Qy 92 GluAlaThrSerTrpProAsnLeuCysLeuGlnAsnSerGlnTrpLeuLeuAspGln 111
Db 256 GAAGCCACTCTTAACCTTAATTTGTGCTTCAAGAACTAAGATGTGCTCTTAATATCA 315
Qy 112 HisMetLeuLysValHisIleTrpProLysPheGlyValSerGluAspCysLeuTrpLeuAsn 131
Db 316 CATATGCTCAAGTGCATTAACCCGAATTCGAGATGTCAAGAACTGCTCTACCTGAAC 375
Qy 132 IleTrpAlaProAlaHisAlaAspThrGlySerLysLeuProValIleuValTrpPhePro 151
Db 376 ATCTATGCGCTGCTCCACAGCGGATTAACAGGCTCAAGCTCCCGCTTGTGTGTGCTCCA 435
Qy 152 GlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAlaTrp 171
Db 436 GAGGTGCTTCAAGACTGCTGCTCAAGCTCCATCTTGAATGGTCCGCTGCTGCTCCAT 495
Qy 172 GluAspValLeuValValValGlnTrpArgLeuGlyIlePheGlyPhePheThrThr 191
Db 496 GAGGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555

```

```

RESULT 6
US-10-451-168-40
; Sequence 40, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GPO0039
; CURRENT APPLICATION NUMBER: US/10/451,168

```


QY 192 TPAAPGNIHSAIAProGlyAAsnTPAlaPheLysAspGluValAlaIleuSerTrp 211
DB 556 TGGATCAGCATGCTCCGGGGAACCTGGCCCTTCAAGACAGAGTGGCTGCTCTGG 615
QY 212 ValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu 231
DB 616 GTCCAGAGAAACATCGAGTCTTCGGTGGGAGCCCAAGCTCTGAGCAATCTTGGGAG 675
QY 232 SerAlaGlyAlaIleSerValSerSerIleIleuSerProMetAlaLysGlyLeuPhe 251
DB 676 TCCGCGGAGACCATTAATGCTTTCTAGCTTTATATGCTCCATGGCCAAAGGCTTATTC 735
QY 252 HisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAspTyr 271
DB 736 CACAAAGCCATCATGAGAGTGGGGTGGCATATCCCTTACCTGAGGCCCATGATTAT 795
QY 272 GlnLysSerGluPheLeuGluValAlaIleHisPheCysGlyValAsnAsnLysAspSer 291
DB 796 GAGAAAGTGAAGACCTGAGGTTGGACATTTCTGGGTAAACAATGCTCAGACTCT 855
QY 292 GlnAlaIleuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrIleuSerGln 311
DB 856 GAGGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGCTGACCTCAGCCAG 915
QY 312 LysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 331
DB 916 AAAACAAGCTTTCCTGAGAGTGTGATGGTCTTCTTCTTAAGCCTCATAGT 975
QY 332 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnIleGlu 351
DB 976 CTATTGCTCAGAAACATTTAAAGCAATTCCTTCATCATGAGTCAATTAACACAGAG 1035
QY 352 CysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLysSer 371
DB 1036 TGTGGCTTCTCTGCTGCTATGAAGAGGCTCCAGATCTCAGTGGCTCCAGAACAGTCC 1095
QY 372 LeuAlaIleuHisIleuIleGlnAsnIleLeuHisIleProGlnTyrIleuHisLeuVal 391
DB 1096 CTTCGCCCTTCATGATCAAAACATCCGACATCCGCCCTCAATATTTGGACCTTGTG 1155
QY 392 AlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeuAsp 411
DB 1156 GCTAATGAAATACCTTCATGAGCAAGCACTCCGATCGAAATCCGAGACAGCTTCTGAGC 1215
QY 412 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAsp 431
DB 1216 TTGCTTGAAGATGCTGCTTCTTGTGCTCCCTGCACTGATCAGAGCTCATACAGAGAT 1275
QY 432 AlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAspThr 451
DB 1276 GCTGGTGACCTGCTTACTTCTATAGATTGAGCACCGGCTCAGTGTCTTGAAGACAGC 1335
QY 452 LysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGlyAla 471
DB 1336 AAGCAGCTTTTGTCAAGCCGAGCAAGCTGATGAAGTCCGCTTGTGTTGGTGGTGGCC 1395
QY 472 PheLeuLysGlyAspIleValMetPheGluGluValArgGluGluGluLysLeuSer 491
DB 1396 TTCTTGAGGGGAGCATTTGATGTTGAGAGAACCCAGAGAGGAGAGATTACTAGC 1455
QY 492 ArgLysMetMetLysTyrTyrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 511
DB 1456 CGGAAGATGATGAAATACCTGGGCTACCTTGGCTCGAACCGGGAATCTTAATGGAGACAGC 1515
QY 512 LeuSerLeuTyrProAlaTyrAsnLeuThrGluGlnTyrLeuGluLeuAspLeuAsnMet 531
DB 1516 CTGCTCTGTGGCAGCTTATATCTGATGAGCAGTACCTCCAGCTGAGACTTGAACATG 1575
QY 532 SerLeuGlyGlnArgLeuLysGluProArgValAspPheThrProSerThrIleProLeu 551
DB 1576 AGCTCGAGACAGACCTCAAAAGAACCGGGGTGATTTTGGACCAAGCAATCCCTGG 1635
QY 552 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 571

DB 1636 ATCTGCTGCTCCGAGACATGCTCCAGAGTCTCTTCTTCAATTTCTCTCTC 1695
QY 572 LeuGlnProPhePhePheCysAlaPro 581
DB 1696 CTCAGACCTTTCTTTTCTTTTGTGCTCT 1725
RESULT 7
US-10-433-256-23
; Sequence 23, Application US/10433256
; Publication No. US20040081980A1
GENERAL INFORMATION:
; APPLICANT: SANTANWALA, Madhundan M.; YAO, Monique G.
; APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
; APPLICANT: ARVIZO, Chandra S.; RING, Huijun Z.
; APPLICANT: LEE, Ernestine A.; DING, Li
; APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
; APPLICANT: YUE, Henry; TRIBOULEY, Catherine M.
; APPLICANT: LU, Dyrung Alina M.; LAU, Preeti G.
; APPLICANT: WARREN, Bridget A.; YANG, Junming
; APPLICANT: GHAMLA, Narinder K.; NGUYEN, Daniel B.
; APPLICANT: GANDHI, Ameena R.; LU, Yan
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0313 USN
; CURRENT APPLICATION NUMBER: US/10/433,256
; PRIOR APPLICATION NUMBER: PCT/US01/47429
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/254,308
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/256,189
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/257,713
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/262,706
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/266,020
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ. ID NOS: 26
; SOFTWARE: PERL Program
; SEQ ID NO 23
; LENGTH: 2232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 6538080CB1
US-10-433-256-23
Alignment Scores:
Pred. No.: 0 Length: 2232
Score: 2900.00 Matches: 548
Percent Similarity: 99.82% Conservative: 2
Beef Local Similarity: 99.46% Mismatches: 1
Query Match: 94.19% Indels: 0
Gaps: 0
US-10-023-515-2 (1-581) x US-10-433-256-23 (1-2232)
QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTyrPileGlnGlyLysGln 50
DB 274 GGGCCCTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGGATTCAGGGCAAGCA 333
QY 51 ValThrValLeuGlySerProValProAlaAsnValPheLeuGlyValProPheAla 70
DB 334 GTTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTCTCGAGAGTCCCTTGTGCTGCT 393
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrProAsnLeu 90
DB 394 CCCCCTGGAGTCCCTGATTTAAGAACCCGAGACGCTGATGCGCCCTGGGATTAACCTTG 453
QY 91 ArgGluAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTyrPheLeuLeuAsp 110

7 1125100
102e OK


```

Db 454 CGAAGACCACTCTACCTAATTGTCCTCCAGAACTCAGAGGCTGCTTACAT 513
Qy 111 GlnHsmELeuLyValHsTyProLyRheGlyValSerGluAspCyLeuTyLeu 130
Db 514 CAACGATGCTCAAGGTGATTAACCGAAATTCGAGGTCCAGAAAGACGCTCTACCG 573
Qy 131 AsnIeTyAlaProAlaHsIAlaAspThrGlySerLeuLeuProValLeuValTrpPhe 150
Db 574 AACATCTATGGCTCGCCACGCCCATACAGGCTCCAACTCCCTGTTGGTGTTC 633
Qy 151 ProGlyGlyAlaPheLyThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
Db 634 CCAGAGGAGGCTTCAAGACTGCTCAAGCTCATCTTGTGATGGGACCGCTGCTG 693
Qy 171 TyrgIuAspValLeuValValValGlnTyArgLeuGlyIlePheGlyPheThr 190
Db 694 TATGAGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 753
Qy 191 ThrTrpAspGlnHsIAlaProGlyAsnTrpAlaPheLyAspGlyValAlaAlaLeuSer 210
Db 754 ACATGGGATCAGCATGCTCTCGGGAACTGGGCTTCAGAGACCAAGTGTCTGTCT 813
Qy 211 TrpValGlnLyAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
Db 814 TGGGTCCAGAAAGACATCGAGTCTTTCGGTGGGACCCAGCTGTGTGACATCTTGGC 873
Qy 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLyGlyLeu 250
Db 874 GAGTCCGGGGGAGCATAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 933
Qy 251 PheHsLyAlaIleMetGluSerGlyValAlaIleIleProTyLeuGluAlaHsAsp 270
Db 934 TTCACAAAGCATATGAGAGGTGGGGGAGGCTCATCTCTTACCTGAGGCTCATGAT 993
Qy 271 TyrgIuLySerGluAspLeuGlnValAlaAlaHsPheCyGlyAsnAlaIleSerAsp 290
Db 994 TATGAGAAAGTGAAGACCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1053
Qy 291 SerGluAlaLeuLeuArgCyLeuArgThrLyPProSerLyGluLeuLeuThrLeuSer 310
Db 1054 TCTGAGGCTCTGCTGAGGTGCTGCTGAGCAAAACCTCCAGAGAGCTGTGACCTCAGC 1113
Qy 311 GlnLyThrLySerPheThrArgValAlaAspGlyAlaPhePheProAsnGluProLeu 330
Db 1114 CAGAAACAAAGTCTTCACTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1173
Qy 331 AspLeuLeuSerGlnLyAlaPheLyAlaIleProSerIleIleGlyValAlaAsnHs 350
Db 1174 GATCATTTGCTCAGAAAGCATTTAAAGCAATTCCTTCATCATATCGAGTCAATTAAC 1233
Qy 351 GlnCyGlyPheLeuLeuProMetLyGluAlaProGluIleLeuSerGlySerAsnLy 370
Db 1234 GAGTGTGCTTCTCTGCTCTGTGAAGAGGCTCTGAGGCTCTCAGTGTCTCAACAG 1293
Qy 371 SerLeuAlaLeuHsIleuIleGlnAsnIleLeuHsIleIleProProGluIleuHsIleu 390
Db 1294 TCCCTTGGCTCTCATCTGATAAAACATCTGCAATCCGCTCAGATTTGCACTT 1353
Qy 391 ValAlaAsnGlyTyPheHsAspLyHsSerLeuThrGluIleArgAspSerLeuLeu 410
Db 1354 GTGGCTATGAAATATCTTCATGACAAAGCACTCCCTGACATGAATCCGAGACAGTCTTCG 1413
Qy 411 AspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyHsArg 430
Db 1414 GACTTGTCTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1473
Qy 431 AspAlaGlyAlaProValTyPheTyrgIuPheArgHsArgProGlnCyPheHsIleu 450
Db 1474 GATGCTGGGACCTGCTACTTCTATGAGTTCGAGCCGCTCAGGCTTCAAGAC 1533
Qy 451 ThrLyProAlaPheValLyAlaAspHsAlaAspGlyValArgPheValPheGlyGly 470

```

```

Db 1534 ACGAAGCCGCTTTTGTCAAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGTGTGT 1593
Qy 471 AlaPheLeuLyGlyAspIleValMetPheGluGlyAlaThrGluGluGlyLeuLeu 490
Db 1594 GCTTCTCGAAGGGGAGCAATTTGTAATGTTTCAAGAGGCCACGAGAGAGAAATTAATCTG 1653
Qy 491 SerArgLyMetMetLyTyTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 510
Db 1654 AGCCGGAAGATGATGAATATCTAGGCTACCTTTGTCTCAACCGGAAATCTTAATGGAAAC 1713
Qy 511 AspLeuSerLeuTrpProAlaTyAsnLeuThrGluGlnTyLeuGlnLeuAspLeuAsn 530
Db 1714 GACCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1773
Qy 531 MetSerLeuGlyGlnArgLeuLyGluProArgValaAspPheTrpThrIlePro 550
Db 1774 ATGAGCTTCGACAGAGACTCAAAAGACCGGGGTGAGTGTGTGTGTGTGTGTGTGTGT 1833
Qy 551 LeuIleLeuSerAlaSerAspMetLeuHsSerProLeuSerSerLeuThrPheLeuSer 570
Db 1834 CTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1893
Qy 571 LeuLeuGlnProPhePhePhePheCysAlaPro 581
Db 1894 CTCTCCAGCCTTTCTTTTCTTTTGTGTGTCTCT 1926

RESULT 8
US-10-104-047-249
; Sequence 249, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-249

Alignment Scores:
Pred. No.: 0 Length: 2092
Score: 2877.00 Matches: 551
Percent Similarity: 98.40% Conservative: 2
Best Local Similarity: 98.04% Mismatches: 2
Query Match: 93.44% Indels: 7
DB: 17 Gaps: 2

US-10-023-515-2 (1-581) x US-10-104-047-249 (1-2092)
Qy 24 HsAspGlnTrpGlyLyThrGlyProSerAlaGlyGlyProGlnArgAsnThrArgLeu 43
Db 126 CACTTAACTGG-----GGGCTTCTGTGTGAAGGCCACAGAGAAACCAAGGCTG 176
Qy 44 GlyTrpIleGlnGlyLyGlnValThrValLeuGlySerProValProValAsnValPhe 63
Db 177 GGAATGATTCAGGGCAAGCATGCTGTGCTGGAGCCCTGTGCTGTGAACGTGTTC 236
Qy 64 LeuGlyAlaProPheAlaAlaProProLeuGlySerLeuArgPheThrAsnProGlnPro 83
Db 237 CTCGAGTCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 296
Qy 84 AlaSerProTrpAspAsnLeuArgGlyAlaThrSerTyProAsnLeu-----Cy 100
Db 297 GCATCGCCTCGGATTAACCTTCGAGAAACCACTTCCTTAATTTGTAAAGACAGGTG 356
Qy 100 sLeuGlnAsnSerGluTrpLeuLeuAspGlnHsMetLeuLyValHsTyProLy 120

```

Db 357 CCTCCAGAACTCAGAGTGGCTGCTTAGATCAACATGCTCAAGTGCATTACCGAA 416
Qy 120 sPhag1YValSerGluuPcySerGlyLeuHs1eTyAlaProAlaHs1a1aPth 140
Db 417 ATTGGAGGTGTCAGAGAGCTGCTCTACCTGAAATATATGCGCTCCACCGCATAC 476
Qy 140 rG1ySerLysLeuProValLeuVal1TPheProG1yG1yAlaPheLysThG1ySerL 160
Db 477 AGGCTCCAGAGCTCCCGCTTGGTGTGCTCCAGAGAGTCCCTTCAAGCTGACTCAGC 536
Qy 160 aSerL1ePheAapG1ySerAlaLeuAlaAlaTyG1yAapValLeuVal1yAlaG1 180
Db 537 CTCATCTTTGATGGGTCCCGCTGGCTCTATAGAGACGTGCTGGTGTGTCTGTCCA 596
Qy 180 nTyArGLeuG1yL1ePheG1yPhePheThrThrTPAapG1nHs1a1aProG1yAant 200
Db 597 GTACCGGCTAGAAATATTGGTTCTTCCACCAATGGAGATGACATGCTCCGGGGAATG 656
Qy 200 pAlaPheLysAapG1nVal1aAlaLeuSerTPValG1nLysAant1eG1yPhePheG1 220
Db 657 GGCCTTCAAGAGCAAGGTGGCTGCTGTCTGGTCCAGAAAGACATGAGTTCTTCCG 716
Qy 220 yG1yAapProSerSerValThrL1ePheG1yG1ySerAlaG1yAlaL1eSerValSerSe 240
Db 717 TGGGAGACCCAGCTGTGACCATTTTGGCGAGTCCGGGAGCATATAGTTTCTAG 776
Qy 240 rLeuL1eLeuSerProMetAlaLysG1yLeuPheHs1yLysAlaL1eMetG1ySerG1yVa 260
Db 777 TCTTATACGTGTCTCCATGGCCAAAGCTTATTCCAAAGCCATCATAGAGTGGGT 836
Qy 260 lAlaL1eL1eProTyrLeuGluAlaHs1aPtyrG1yLysSerG1yAapLeuG1nVal1Va 280
Db 837 GGCCATCATCCCTTACCTGAGAGCCCATATATAGAGAGTGAAGTCACTGCACTGCTG 896
Qy 280 lAlaHs1PheCyG1yAapAapAlaSerAapSerG1nAlaLeuAaTyG1yLeuAryTh 300
Db 897 TGCACATTTCTGTGTGTAACATGCTCAAGCTTGAAGCCCTGAGAGTGTCTTGAAGAC 956
Qy 300 rLyAapProSerLyG1yLeuLeuThrLeuSerG1nLysThyLysSerPheThrAryVal1Va 320
Db 957 AAACCCCTCCAAAGAGTGTGCTGACCTCAGCAAAACAAAGTTTCACTCAGTGTGT 1016
Qy 320 lAapG1yAlaPhePheProAapG1uProLeuAapLeuSerG1nLysAlaPheLysAl 340
Db 1017 TGATGCTGCTTTCTTCTTAAGACCTCTAGATATATGCTCAAGAAACATTTAAAGC 1076
Qy 340 aL1eProSerL1eL1eG1yValAapAapHs1eG1yCyG1yPheLeuLeuProMetLysG1 360
Db 1077 AATTCCTTCCATCATCGAGTCAATAGCCAGAGTGTGCTTCTGCTGCTTATAGAGGA 1136
Qy 360 uAlaProG1nL1eLeuSerG1ySerAapLysSerLeuAlaLeuHs1eLeuL1eG1nAant1 380
Db 1137 GGCCTCGAGATCTCAGTGTGCTCCAAACAGTCCCTTGCCTCATCATCAAAACAT 1196
Qy 380 eLeuHs1eL1eProProG1nTyLeuHs1eLeuVal1AlaAapG1uTyPheHs1aPheLysH 400
Db 1197 CCTGCAATCCCGCTCAGATTGCTGACCTTGTGCTTAATGAATACTTCCATGCAAGCA 1256
Qy 400 sSerLeuThrG1nL1eAryAapSerLysLeuAapLeuG1yAapVal1PhePheVal1Va 420
Db 1257 CTCCTCACTGAATCCGAGACAGTCTTCTGGACTTCTGAGAGATGTGTTCTTGTGCT 1316
Qy 420 lProAlaLeuL1ePheThrAlaAryTyHs1aPheAapAlaG1yAlaProVal1TyPheTyG1 440
Db 1317 CCTGCACTGATCAAGCTGATATCAAGAGATGCTGAGACCTGTCTCATCTTATGCA 1376
Qy 440 uPheArG1Hs1aPtyrG1nCyPheG1yAapThrLysProAlaPheVal1yAlaPth 460
Db 1377 GTTTCGCAACCGGCTCAGTCTTGAAGACAGAAACCGGCTTTGTCAAAACCCACCA 1436
Qy 460 sAlaAapG1nVal1AryPheVal1PheG1yAlaPheLeuLysG1yAapL1eVal1ePhePh 480

Db 1437 CGCTGATGAAGTCCGCTTGTGTGCTGTGCTCTTCTTGAAGGGGAGACATTGTANTGT 1496
Qy 480 eG1yG1yAlaThrG1nG1nG1nLysLeuLeuSerAryLysMetLysTyTTpAlaTh 500
Db 1497 CGAAGAGACGACGAGAGAGAGAGTCTAGACCGAAGATGATGAATATCTGGCTAC 1556
Qy 500 rPheAlaAryThrG1yAapProAapG1yAapAapLeuSerLeuThrPProAlaTyAantLe 520
Db 1557 CTTTGTCTGAACCGGAGATCTTAATGGAGACGACTGTCTGTGGCAGCTTATANTCT 1616
Qy 520 uThrG1nG1nTyLeuG1nLysAapLeuAapMetSerLeuG1yG1nAryLeuLysG1uP 540
Db 1617 GACTGAGAGTACTGCTGACGCTGAGCTTGAACATAGCCTCGACAGACACTCAAAAGAAC 1676
Qy 540 oAryValaPhePheThrThrSerThrL1eProLeuL1eLeuSerAlaSerAapMetLeuH 560
Db 1677 GCGGTGAGATTGTTGACACAGACATCCCTGATCTGTCTGCTCCACATGCTTCA 1736
Qy 560 sSerProLeuSerSerLeuThrPheLeuSerLeuLeuG1nProPhePhePheCyAl 580
Db 1737 CAGTCTCTTCTTCTTAACTTCTCTCTCTCTCCACGCTTTCTTTCTTGTGTC 1796
Qy 580 aPro 581
Db 1797 TCCT 1800

RESULT 9

US-10-451-168-41
; Sequence 41, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; PRIOR FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-41

Alignment Scores:

Pred. No.: 6.97e-303 Length: 1746
Score: 2607.50 Matches: 499
Percent Similarity: 92.67% Conservative: 7
Best Local Similarity: 91.39% Mismatches: 25
Query Match: 84.69% Indels: 15
DB: 18 Gaps: 3

US-10-023-515-2 (1-581) x US-10-451-168-41 (1-1746)

OY		12	TPrpGlnPhePheLeuIleuLeuInProIeuLeuGlyHisArgInTrpGlyLysThrGlu	31
Dd		40	TGGGTACTTGGGTCTTGCAAGCCCC-----ACCAAAGG	75
OY		32	ProSerAlaGluGlyProGlnAraGanThraArgLeuGlyTrpIleGlnGlyLysGlnVal	51
Dd		76	CCTTCTGGCGAAAGGCCCAcAGAAGAACCCAGGCTGGATGGATTCAAGGCACAAGATC	135
OY		52	ThrValIleuGlySerProValProValAsnValPheLeuGlyValProPheAlaIlaPro	71
Dd		136	ACTGTGCTGGGAAGACCCTGTGCTGTGAACGTGTTCTCGAGAGTCCTTTGTGCTGCTCC	195
OY		72	ProIeuGlySerLeuAraPherThraSnProGlnProAlaSerProTrpAraSnaIleuAraG	91
Dd		196	CCGCTGGGATCCCTCGCATTTACGAACCCGACGCTGATCGCCCTGGGAATACTTGGGA	255
OY		92	GluAlaTrpSerTyProAsnIleuCyLeuGlnAsnSerGluTrpLeuIleuAspGln	111
Dd		256	GAAGCACCTCTTAACCTTAATTGTGGCTCCAGAAGCTAGAGTGGCTCTTAGATADA	315
OY		112	HisMetIleuLysValHisTyProLysPheGlyValSerGluAspCyLeuTyTrpLeuAsn	131
Dd		316	CATATGCTCAAGGTCAATTACCCGAATTCGAGGTCAAGAAAGTCGCTCTACCTGAAC	375
OY		132	IleTyTrAlaProAlaHisAlaAspThrGlySerLysLeuProValIleuValTrpPhePro	151
Dd		376	ATCATATGGCCCTGCCACGCGATCAAGGCTCCAAGCTCCCCTGTTGTGTGGTTCCCA	435
OY		152	GlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIlaTyR	171
Dd		436	GAAGGTGCTTCAGACAGTGGCTCAAGCTCCACATTGTAGTGGTTCGCGCTGACCTAT	495
OY		172	GluAspValLeuValValValGlnTyraArgLeuGlyIlePheGlyPhePheThrThr	191
Dd		496	GAGAGCGTGGCTGGTGTGTGTGTCCAGTACCGGCTAAGGAATATTGGTTCTTCAACCA	555
OY		192	TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaIleuSerTrp	211
Dd		556	TGGGATCAGCATGCTCCGCGGGAACCGGCGCTTCAAGGACAGAGTGGCTGCTGCTCGG	615
OY		212	ValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu	231
Dd		616	GTCCAGAAGAACATGAGATCTTCTCGTGGGAGCCCACTCTGTGACATTTTGGCCAG	675
OY		232	SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPrometAlaLysGlyLeuPhe	251
Dd		676	TCCGCGGAGCCATAGTGTCTTCAAGTCTTATACGTCTTCCCAAGGCGCAAGGCTTATTC	735
OY		252	HisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAspTyR	271
Dd		736	CACAAAGCCATCATGAGAGAGTGGGGTGGCCATCATCCTTACCTGGAAGGCCCATGATTAT	795
OY		272	GluLysSerGluAspLeuGlnValValAlaHisPheCySGlyAsnAsnAlaSerAspSer	291
Dd		796	GAGAAAGATGAGACCTTCAGGTGTTTCACATTTCTGTGGTAACTAAGCGTCAAGACTCT	855
OY		292	GluAlaLeuLeuAraCyLeuAraGthrTyrsProSerLysGluLeuLeuThrLeuSerGln	311
Dd		856	GAAGGCCCTGCTGAAGTGGCTGAAGCAAAACCTTCCAAGAGAGTGAACCTTACCGCAG	915
OY		312	LysThrLysSerPheThrArGValValAlaSpGlyAlaPhePheProAsnGluProLeuAsp	331
Dd		916	AAAACAAGTCTTTCACCTCGAGTGGTTGATGTTCTTCTTCCATAGAACCTCTAGAT	975
OY		332	LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHisGlu	351
Dd		976	CTATTGTCTCAAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGAGTCAATPAACCAAG	1035
OY		352	CysGlyPheLeuLeuProMetLysGlu-----AlaProGluIleLeuSerGlySerAsn	369
Dd		1036	TGTGGCTTCCGCTGCTCATGTGTAAAGAAATTCGTGGTTCATACCTGCCATCTCCCTCAAC	1095

QY	370	YysrseuLeuAlaLeu-----HisLeuIegiAnanilleuHilLeuPro	384

Db	1096	CGTGTGACGCTTTGGCTTCAACAGCTGGGCACTTTCCACAGAAACATCAGCAATCCG	1155S
QY	385	ProGlnTyLeuHilAleuValAlaAngluTyPheHisApLySHiSserLeuThrglu	404
Db	1156	CCTGAGTATTTGCACCTTGTGGCTTAATGATATCTTCATGACAAAGCACTCCCTGACTGA	1215S
QY	405	IleArgApSerLeuLeuApLeuLeuGluYAspValPhePheValValProAlaLeuIle	424
Db	1216	ATCCAGACACACTCTTCGACCTTCCTGGAGATGTGTTCTTTGGTCCCTGCACCTGATC	1275S
QY	425	ThrAlaArgTyTrHisaArgAspAlaGluAlaProValTyPheTyrgluPheArgHisArg	444
Db	1276	ACAGCTCAATATCAACAGAGATGCTGTGCACCTGTCTATCTTATAGATTTCGGCACCGG	1335S
QY	445	ProGlnCyPheGluApSerThyLysProAlaPheValIysAlaAspHisAlaAspGluVal	464
Db	1336	CCTCAGTCTCTTGAAGACACGAAACCGACTTTTGTCAAAGCCGACACAGCTGATGAAATC	1395S
QY	465	ArgPheValPheGluGlyAlaPheLeuLysGluYAspIleValMetPheGluGlyAlaThr	484
Db	1396	CGCTTTGTGTTGGGTGGTCCCTTCCTGAAGGGGACATGTGTTATGTTCGAAGAGCACCG	1455S
QY	485	GluGluGluLysLeuLeuSerArgLysMetMetLysTyTrTPAlaThrPheAlaArgThr	504
Db	1456	GAGAGGAGGAAGTTATCTGAGCCGGAAGATGATGAATAATCTGGGGTACTTGTGCTCGAAC	1515S
QY	505	GlyAsnProAngIYAsnApLeuSerLeuTyTrProAlaTyraLeuLeuThrgluGlnTyr	524
Db	1516	GGGAATCTTAATGGAAACGACTTCTCTGTGGCCAGCTTTAATCTGACTGAGCAATAC	1575S
QY	525	LeuGlnLeuApLeuAsnMetSerLeuGluYAlaArgLeuLysGluProArgValAspPhe	544
Db	1576	CTCCAGCTGAGCTTGACATATGAGCTCGGACAGAGACTCAAAAGAACCGGAGAGATGTG	1635S
QY	545	TrpThrSerThrIlePro	550
Db	1636	TGGGTGACGGGGTATCCT	1653
RESULT 10			
US-10-094-749-736			
; Sequence 736, Application US/10094749			
; Publication No. US20030219741A1			
; GENERAL INFORMATION:			
; APPLICANT: ISOGAI, TAKAO			
; APPLICANT: SUGIYAMA, TOMOYASU			
; APPLICANT: OTSUKI, TETSUJI			
; APPLICANT: WAKAMATSU, AI			
; APPLICANT: SATO, HIROYUKI			
; APPLICANT: ISHII, SHIZUKO			
; APPLICANT: YAMAMOTO, JUN-ICHI			
; APPLICANT: ISONO, YUUKO			
; APPLICANT: HIO, YURI			
; APPLICANT: OTSUKA, KAORU			
; APPLICANT: NAGAI, KEIICHI			
; APPLICANT: IRIE, RYOTARO			
; APPLICANT: TAMECHIKA, ICHIRO			
; APPLICANT: SEKI, NAOHICO			
; APPLICANT: YOSHIKAWA, TSUTOMU			
; APPLICANT: OTSUKA, MOTOFUKU			
; APPLICANT: NAGAHARI, KENJI			
; APPLICANT: MASUHO, YASUHIKO			
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA			
; FILE REFERENCE: 084335/0160			
; CURRENT APPLICATION NUMBER: US/10/094, 749			
; CURRENT FILING DATE: 2002-03-12			
; PRIOR APPLICATION NUMBER: 60/350, 435			
; PRIOR FILING DATE: 2002-01-24			
; PRIOR APPLICATION NUMBER: JP 2001-328381			
; PRIOR FILING DATE: 2001-09-14			
; NUMBER OF SEQ ID NOS: 3381			
; SOFTWARE: PatentIn Ver. 2.1			

```

; SEQ ID NO 736
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-736

Alignment Scores:
Pred. No.: 1,296-302 Length: 1962
Score: 2606.00 Matches: 503
Percent Similarity: 88.77% Conservative: 3
Best Local Similarity: 88.25% Mismatches: 6
Query Match: 84.64% Indels: 58
DB: 17 Gaps: 2

US-10-023-515-2 (1-581) x US-10-094-749-736 (1-1962)

QY 12 TTPCySPhePheLeuLleLeuGlnProLeuLeuGlnHisArgGlnTrpGlyLeuThrGly 31
D 189 TGGGGCTATCTGGGCTCTTGCAGACCC-----ACCAAGGG 22
QY 32 ProSerAlaGlnGlyProGlnArgAspThrArgLeuGlyTrpLleGlnGlyLeuVal 51
D 225 CTTCTCTGTGAAGGGCCACAGAGAACCCAGGCTGGAGTGAATTCAGGCGACCAATC 284
QY 52 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAlaPro 71
D 285 ACTGTGTGGGAAACCCCTGTCCTGTGAACGTGTTCTCCGAGTCCCTTTGCTGCTCC 344
QY 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArg 91
D 345 CCGTGGGATTCCTGGCCAGATTATCGAACCCGAGCCTGCATCCCTGGAGTAACCTGCCA 404
QY 92 GluAlaThrSerTrpProAsnLeuCyLeuGlnAsnSerGluTrpLeuLeuAspGln 111
D 405 GAAGCCACCTCCATCCCTAAATTGTGCTCCAGAACTCAGAGTGCCTCTTATGATCAA 464
QY 112 HisMetLeuArgValHisTrpProLysPheGlyValSerGluAspCysLeuTrpLeuAsn 131
D 465 CACATGTCTCAAGGTCATTACCCGAAATTCGAGTGCACAAACATGCTCTACCTGAAC 524
QY 132 IleTrpAlaProAlaHisAlaAspThrGlySerTrpLeuProValLeuValTrpPhePro 151
D 525 ATCTATGGCGCTCCGCCACGCCGATACAGGCTCCAAAGTCCCTCTTGGTGTGTTCCCA 584
QY 152 GlyGlyValAlaPheLysThrGlySerAlaSerTrpPheAspGlySerAlaLeuAlaValTrp 171
D 585 GAGAGTGGCTTCAAGACTGCTCAGGCTCATCTTGATGGTCCGCCCTGGCTGCTAT 644
QY 172 GluAspValLeuValValValGlnTrpArgLeuGlyTrpLlePheGlyPhePheThrTrp 191
D 645 GAGGAGCTGCTGTTGTGTGTCGCCAGTACCGGCTAGAAATTTGTTCTTACACAA 704
QY 192 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrp 211
D 705 TGGGATCACACAGTCTCCGGGAACTGGGCTTCAAGACAAGGTGCTCTCTGTCCTGG 764
QY 212 ValGlnLysAsnLleGluPhePheGlyGlyAspProSerSerValThrLlePheGlyGlu 231
D 765 GTCCAGAAAGACATCGAGTTCCTCGGTGGAGACCCAGCTCTGTAGCAATCTTTGGCGAG 824
QY 232 SerAlaGlyAlaTrlSerValSerSerLeuLleLeuSerProMetAlaLysGlyLeuPhe 251
D 825 TCCGGGGGAGCATTAAGTGTTCAGTCTTATACGTCTTATGCCATGGCCAAAGCTTATTC 884
QY 252 HisLysAlaAlaMetGluSerLysLysAlaLleLleProTrpLeuGluAlaHisAspTrp 271
D 885 CACAAAGCATATGAGAGTGGGGGCGCATCATCCCTTACCTGGAAGCCCATATTAAT 944
QY 272 GluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 291
D 945 GAGAAAGATGAGGACCTGAGGTGTTGCACATTTCTGTGTTACAAATGCCATCAACTCT 1004
QY 292 GluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSerGln 311

```

Db	1005	GAGGCGCTGCTGAGTGCCTGAGACAAACCTCCAGAGACTGCTGACCTTCAGCCAG	106
QY	312	LYeThriYSeRphethrArGValValaApGValAlaPhePheProAvnGluProLeuAap	331
Db	1065	AAACCAAGCTTTCCTGAGAGGTGATGGGTCTTCTTCCTAATGAGCTCTAGAT	1122
QY	332	LeuLeuSerGlnYValaPheYValaAlleProSerTlelleGlyValaVnAAnHlsglu	351
Db	1125	CTATTGCTCAGAAACATTTAAAGCAATTCCTTCATTTACGAGTCAATTAACACGAG	1184
QY	352	CYagilyPheLeuLeuProMetLySGlAlaProGluileLeuSeSGlSeSaAnlySeS	371
Db	1185	TGAGGCTCTCTGCTGCTGCTAAGAGAGGCTCCTGAGACCTCAGTGGCTCCAAAGTCC	1244
QY	372	LeuAlaLeuHlVleuTleGlnAnHlleuHlslleProProGlnTyrLeuHlVleuVal	391
Db	1245	CTTGCCCTCATCTGTATACAAACATCCGACATCCGCGCTCAGTATTGCACTTGCTG	1304
QY	392	AlaAAnGluTyrPheHlSaBrLyHlSeSerLeuTnGluHlleAaAaSeSerLeuLeuAap	411
Db	1305	GCTAAAGATCTTCCTCATACAGCACTCCCTGACTGAAATCCAGACAGTCTTCTGGAC	1364
QY	412	LeuLeuGlyAaPValaPhePheValValProAlaLeuHlLeThraAlaGlyTnHlAaP	431
Db	1365	TTCGCTTGAGATGTGTCTTGTGCTGCTGACGTGATCAGCTCATATACAGAG---	1421
QY	432	AlaGlyAlaProValTyrPheTyrGluPheArgHlSaBrProGlnCyPheGluAaPThr	451
Db	1421	-----	1421
QY	452	LyPProAlaPheVallyAlaAaPnHlSaAlaAaPgluValArgPheValPheGlyGlyAla	471
Db	1421	-----	1421
QY	472	PheLeuLyVgGlyAaPrlleValMeSePheGluGlyAlaThrgluGluGluYVleuLeuSer	491
Db	1422	-----GAAAGAACCCAGGAGGAGGAGGAGAGTAACTGAGC	1454
QY	492	ArgLyMeMetLeuYTrTPrAlaThrPheAlaArgHlGlyAaPProAAnGlyAaAaP	511
Db	1455	CGGAAGATGATGAATACTGGGCTAACCCTTGCTGAACCGGGAATCCTAATGGGAAACGAC	1514
QY	512	LeuSerLeuTPrProAlaTyrAaAnLeuThrgluGlnTyrLeuGlnLeuAaPleAaAaMet	531
Db	1515	CTGTCTGTGGCCAGCTTATATCTAGCTGAGAGTACTCCAGCTGAGACTTGAAACATG	1574
QY	532	SeerLeuGlyGlnAaGleuLyVgGluProArgValAaPheTPrThrSerThrlleProLeu	551
Db	1575	AGCCTCGACACGAACTCCAAAGAACCCGGGTGGATTTTGGACCGACACATCCCCCTG	1634
QY	552	lleuSeSerAlaSeSaPmetLeuHlSeSProLeuSeSerleuThrPheLeuSeSerleu	571
Db	1635	ATCCTGTGTGCTCCGACATGCTCCACAGTCCCTTCTTCTTAACCTTCTCTCTCTC	1694
QY	572	LeuGlnProPhePhePhePheCyValaPro	591
Db	1695	CTCCAGCCTTCTTCTTCTTGTGGCTCT	1724
RESULT 11			
US-10-114-270-195			
; Sequence 195, Application US/10114270			
; Publication No. US20040030110A1			
; GENERAL INFORMATION:			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Kekuda, Ramesh			
; APPLICANT: Maller, Charles E.			
; APPLICANT: Maljankar, Uriel M.			
; APPLICANT: Szytek, Kimberly A.			
; APPLICANT: Paturajan, Meera			
; APPLICANT: Litu, Ziaohong			
; APPLICANT: Gusev, Vladimir Y.			
; APPLICANT: Id, Ij			

```

: APPLICANT: Vernet, Corine
: APPLICANT: Zernsen, Bryan D.
: APPLICANT: Gorman, Linda
: APPLICANT: Shenoy, Sureeh G.
: APPLICANT: Pena, Carol B.A.
: APPLICANT: Smithson, Glennda
: APPLICANT: Burgess, Catherine B.
: APPLICANT: Gerlach, Valerie
: APPLICANT: Padigaru, Muralidhara
: APPLICANT: Shinkets, Richard A.
: APPLICANT: Gangoli, Esba A.
: APPLICANT: Taupier Jr., Raymond J.
: APPLICANT: Casman, Stacie J.
: APPLICANT: Ji, Weizhen
: APPLICANT: Anderson, David W.
: APPLICANT: Liete, Mario W.
: APPLICANT: Raetelli, Luca
: APPLICANT: Edinger, Shlomit R.
: APPLICANT: Stone, David J.
: APPLICANT: MacDougall, John R.
: APPLICANT: Rothenberg, Mark B.
: TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
: FILE REFERENCE: 21402-322C
: CURRENT FILING DATE: 2002-11-27
: PRIOR APPLICATION NUMBER: US/10/114,270
: PRIOR FILING DATE: 2001-04-03
: PRIOR APPLICATION NUMBER: 60/281,086
: PRIOR FILING DATE: 2001-04-03
: PRIOR APPLICATION NUMBER: 60/281,136
: PRIOR FILING DATE: 2001-04-03
: PRIOR APPLICATION NUMBER: 60/281,863
: PRIOR FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: 60/281,906
: PRIOR FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: 60/282,020
: PRIOR FILING DATE: 2001-04-06
: PRIOR APPLICATION NUMBER: 60/282,930
: PRIOR FILING DATE: 2001-04-10
: PRIOR APPLICATION NUMBER: 60/282,934
: PRIOR FILING DATE: 2001-04-10
: PRIOR APPLICATION NUMBER: 60/283,512
: PRIOR FILING DATE: 2001-04-12
: PRIOR APPLICATION NUMBER: 60/283,710
: PRIOR FILING DATE: 2001-04-13
: PRIOR APPLICATION NUMBER: 60/284,234
: PRIOR FILING DATE: 2001-04-17
: Remaining Prior Application data removed - See file Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 470
: SEQ ID NO 195
: LENGTH: 1746
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(1744)
: US-10-114-270-195

```

Alignment Scores:

```

Pred. No.: 1.12e-301 Length: 1746
Score: 2597.50 Matches: 497
Percent Similarity: 95.07% Conservative: 4
Best Local Similarity: 94.31% Mismatches: 17
Query Match: 84.36% Indels: 9
DB: 18 Gaps: 3

```

US-10-023-515-2 (1-581) x US-10-114-270-195 (1-1746)

```

QY 31 GILPQSEKALAGLGLYPROGLNARGANTRARGLEUGLYTRPILEGNGLYLVSGLN 50
DB 79 GGGCTTCTGCTGAGAGGGCCACAGAGAACACCAAGCTGGATGATTCAGGGCAAGCA 138
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 139 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAGAACGTGTTCTCTCGAGTCCCTTTCTGTCT 198

```

```

QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnAsnLeu 90
DB 199 CCCCCGCTGGATCCCTCGATTTACGAAACCGAGCTGATGCTCCCTGGGATAACTTG 258
QY 91 ArgGlnAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTrpLeuLeuAsp 110
DB 259 CGAGAAAGCCACTCTTACCTTAATTTGTGCTCCAGAACTCAGATGGCTGCTTAAGT 318
QY 111 GlnHisMetLeuValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeu 130
DB 319 CAACATATGCTCAAGGTCAATTAACCGAAATTCGAGTGTCAAGAGACTGCTTACCTG 378
QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhe 150
DB 379 AAGATCTATGCGCCCTGCCACCGCATACAGGCTCAAGCTCCCGCTGTGGTGTTC 438
QY 151 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
DB 439 CCAGAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAATGGGTCCGCTGCTGCC 498
QY 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
DB 499 TATGAGAGCGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 558
QY 191 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaLeuSer 210
DB 559 ACATGGATCAAGCATGCTCCCGGGAACCTGGCTTCAAGACACAGTGGCTGTGTTC 618
QY 211 TrpValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
DB 619 TGGGTCCAGAAAGAACATGAGTTCGTGGTGGAGCCACAGCTGTGTGATCTTTGGC 678
QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 250
DB 679 GAGTCCGGGGGAGCCATAAGTGTCTTGTCTTATACGTCTCCATGAGCCAAAGCTTA 738
QY 251 PheHisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp 270
DB 739 TTCACAAAGCCATATAGAGAGTGGGTGGCCATCATCTTCACTGAGGCCATGAT 798
QY 271 TyrGlnLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAsp 290
DB 799 TATGAGAAAGATGAGGAGCTGAGGTGTGTGCACTTCTGTGTGTAACAATCGTCAGAC 858
QY 291 SerGlnAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer 310
DB 859 TCTGAGGCCCTGCTGAGGTGCTGAGCAAAACCTCCAAAGAGCTGCTGACCTCAGC 918
QY 311 GlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeu 330
DB 919 CAGAAACAAAGTCTTACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 978
QY 331 AspLeuLeuSerGlnLysAlaPheLysValAlaIleProSerIleIleGlyValAsnAsnHis 350
DB 979 GATCTATTGTCTCAGAAAGCATTTTAAAGCATTCCTTCCATCATGAGTCAATAACAC 1038
QY 351 GlnCysGlyPheLeuLeuProMetLysGlu-----AlaProGluIleLeuSerGlySer 368
DB 1039 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1098
QY 369 AsnLysSerLeuAlaLeu-----HisLeuIleGlnAsnIleLeuHisIle 383
DB 1099 AACCGTATGAGCTTGTGCTTCAACAGCTGGGCACTTTCACAGAGAG-----CATATC 1152
QY 384 ProProGlnTyrLeuHisLeuValAlaAsnGlnTyrPheHisAspLysHisSerLeuThr 403
DB 1153 CCGCTCAAGTATTTGCACTTGTGTGTATGATTAATCTTCCATGACAAACACTCCCTGACT 1212
QY 404 GlnIleArgAspSerLeuLeuAspLeuLeuGlnLysAspValPhePheValValProAlaLeu 423
DB 1213 GAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1272

```

```
QY 424 ILeThraLaArgTYrHieARgAspAlaGlyValaProValTYrPheTYrGluPheArgHis 443
Db 1273 ATCAAGCTCGATATTCACAGAGATGCTGTGCACTGTCTATGAGTTTCGGCAC 1332
QY 444 ArgProGlnCYrPheGluAspThrTYrProAlaPheValTYrValaAspHisAlaAspGlu 463
Db 1333 CGGCTCATGCTTGTGAAGACACAGACGCACTTTGTCAAGCCGACCAAGCTGATGAA 1392
QY 464 ValArgPheValPheGlyGlyAlaPheLeuTYrGlyAspHisAlaMetPheGluGlyAla 483
Db 1393 GTCCGCTTGTGCTTGGTGGTGGTCTTCTGAAAGGGGACATGTATATGTTGAAAGGAGCC 1452
QY 484 ThrGluGluGluTYrLeuLeuSerArgTYrMetMetTYrTYrTrpAlaThrPheAlaArg 503
Db 1453 ACCGAGGAGGAGAACTACTGAGCCGGAAGATGATGAATACTGGCTACTTGTCTCA 1512
QY 504 ThrGlyAsnProArgGlnTYrAsnAspLeuSerLeuTYrProAlaTYrAsnLeuThrGluGln 523
Db 1513 ACCGGAAATCTTAATGGAACGACCTGTCTGTGGCCAGCTTATATCTGACTGAGCAG 1572
QY 524 TYrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuTYrGluProArgValaAsp 543
Db 1573 TACCTCAGCTGAGCTTGAACATGAGCTTCGACAGACTCAAGAACCGCGAGAGAT 1632
QY 544 PheTrpThrSerThrIlePro 550
Db 1633 GTGTGGGTGACGGGGTATCTCT 1653

RESULT 12
US-10-381-898-20
; Sequence 20, Application US/10381898
; Publication No. US20040086867A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKY, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIPFIN, Jennifer A.;
; APPLICANT: HAPALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Farrah A.; LAU, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dying Aina M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
; APPLICANT: SANJANWALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
; APPLICANT: CHAWLA, Narinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming;
; APPLICANT: WARREN, Bridget; XU, Yuming; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381,898
; CURRENT FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ. ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7473645CB1
US-10-381-898-20

Alignment Scores:
Pred. No.: 1,516-300 Length: 1857
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 84.07% Indels: 23
DB: 18 Gaps: 1

US-10-023-515-2 (1-581) x US-10-381-898-20 (1-1857)

QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyGln 50
Db 274 GGGCTTCTGCTGTAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAACAA 333
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 70
Db 334 GTCACTGTGCTGGAGAACCTCTGCTGTGTAACGTGTTCTCGAAGTCCCTTTGCTGCT 393
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeu 90
Db 394 CCCCCGCTGGGATCCTTCGATTTACGAAACCCGACGCTGCATCGCCCTGGGATTAACCTG 453
QY 91 ArgGluAlaThrSerTYrProAsnLeuCYrLeuGlnAsnSerGluTrpLeuLeuAsp 110
Db 454 CGAAGAACCACTCTCTACCTTAATTGTGCTCAGAACTCAGAGTGGCTCTTAAGAT 513
QY 111 GluHisMetLeuTYrValHisTYrProLYrPheGlyValSerGluAspCYrLeuTYrLeu 130
Db 514 CAATATATGCTCAAGGATGATTAACCGAAATTCGAGTCTCAGAGACTGCTCTACTG 573
QY 131 AsnIleTYrAlaProAlaHisAlaAspThrArgTYrSerTYrLeuProValLeuValTrpPhe 150
Db 574 AACATCTATCGCTGCCGCCACCGCGATACAGGCTCCAGACTCCCGCTCTTGAGTGTG 633
QY 151 ProGlyGlyAlaPheTYrThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 170
Db 634 CAGAGAGGTGCTTCAAGACTGTGCTCAGCTCCATCTTGAATGGGTCCGCTGTGCTGCC 693
QY 171 TYrGluAspValLeuValValValGlnTYrArgLeuGlyIlePheGlyPhePheThr 190
Db 694 TATGAGGAGCTGCTGTGTTGTGCTGCTGATACCGGCTGAGAAATTTGGTTCTTCAAC 753
QY 191 ThrTPAspGlnHisAlaProGlyAsnTrpAlaPheTYrAspGlnValAlaLeuSer 210
Db 754 ACATGGGATCAGACATGCTCCGGGAACTGGGCTTCAAGAACCAAGTGGCTGCTGCC 813
QY 211 TrpValGlnTYrAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
Db 814 TGGGTCAGAAAGAACATGATGCTTCTGAGTGGGACCCCACTCTGTACCATCTTTGGC 873
QY 231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaTYrGlyLeu 250
Db 874 GAGTCGCGGAGGACATGATGTTTCTAGTCTTAACTGCTTCCATGCGCAAGGCTTA 933
QY 251 PheHisValAlaIleMetGluSerGlyValAlaIleIleProTYrLeuGlnAlaHisAsp 270
Db 934 TTCACAAAGCCATCATGAGAGAGTGGGTGCGCATCATCTTCACTGAGGCCCATGAT 993
QY 271 TYrGluTYrSerGluAspLeuGlnValValAlaHisPheCYrGlyAsnAsnAlaSerAsp 290
Db 994 TATGAGAAAGATGAGGAGACTTCAGTGTGCTGCAATTTCTGTGTAACATGCTGAC 1053
QY 291 SerGluAlaLeuLeuArgCYrLeuArgTYrIleAspSerIleGlyLeuLeuThrLeuSer 310
Db 1054 TCTGAGGCCCTGTGAGGTGCTGAGGACAAACCTTCAAGAGGCTGCTGACCTCAGC 1113
QY 311 GlnTYrThrTYrSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeu 330
```



```

Db      1114 CAGAAAAGTCTTTCACCTGAGTGTGATGGTCTTCTTCTTCAATGAGCCTCTA 1173
QY      |||
QY      331 AapleuSeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAlaSerHis 350
Db      1174 GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTCTCATCATGAGTCAATTAACAC 1233
QY      |||
QY      351 GlnCyGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLys 370
Db      1234 GAGTGTGCTTCTCTGCTGCTATG----- 1257
QY      |||
QY      371 SerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIleuHisLeu 390
Db      1258 -----CACATCCGCTCAAGTATTTGACACTT 1284
QY      |||
QY      391 ValAlaGlnGlyTrpPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeu 410
Db      1285 GTGGCTAATGATATCTTCATGACACAGCACTCCGACTGAATCGAGACAGTCTTCG 1344
QY      |||
QY      411 AapleuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTrpHisArg 430
Db      1345 GACTTGCTTGAGATGTGTTCTTTGTGTGCTGCTGATCATGACTGATATACAG 1404
QY      |||
QY      431 AapAlaGlyAlaProValTrpPheTrpGlyuPheArgHisArgProGlnCyAspPheGluAsp 450
Db      1405 GATGCTGTGACACCTGCTACTTCTATGAGTTTCGGCACCGGCTCAAGTCTTGAGAGAC 1464
QY      |||
QY      451 ThrLysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGly 470
Db      1465 ACGAAGCCAGCTTTGTCAAGGCCACACCGCTGATGAAGTCCGCTTGTGTGCGTGT 1524
QY      |||
QY      471 AlaPheLeuLysGlyAspHisLeuValMetPheGluGlyAlaThrGluGluGluLysLeu 490
Db      1525 GCCTTCTGAAAGGGGACATGTGTATGTTCAGAGGACCCAGGAGAGAGAGAGATTTACG 1584
QY      |||
QY      491 SerArgLysMetMetLysTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsn 510
Db      1585 AGCGGAGATGATGAATATGAGGCTACCTTGTGTCGACCGGAATCTTAATGGGAAC 1644
QY      |||
QY      511 AapleuSerLeuTrpProAlaTrpAsnLeuThrGluGlnIleuGlnLeuAspLeuAsn 530
Db      1645 GACCTGTCTGTGGCCAGCTTATATCTGACTGACAGTACCTCAGCTGACCTTGGAAC 1704
QY      |||
QY      531 MetSerLeuGlyGlnArgLeuLysGluProArgValAspPheTrpThrSerThrIlePro 550
Db      1705 ATGAGCTTCGACACAGACTCAAGAACCCGCGAGAGATGTGTGGGTGACGGGATATCT 1764

```

RESULT 13

US-10-233-933A-3

Sequence 3, Application US/10233933A

Publication No. US2004021471A1

GENERAL INFORMATION:

APPLICANT: Yamashita, Tetsuro

APPLICANT: Miyazaki, Masao

TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER

FILE REFERENCE: SHIG P02US006

CURRENT APPLICATION NUMBER: US/10/233,933A

CURRENT FILING DATE: 2002-09-03

PRIOR APPLICATION NUMBER: JP2002-057908

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 2145

TYPE: DNA

ORGANISM: Felis catus

FEATURE:

NAME/KEY: CDS

LOCATION: (175)..(1803)

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1804)..(2145)

OTHER INFORMATION: n is a, c, g, or t

US-10-233-933A-3

Alignment Scores:

Pred. No.:	2,56-231	Length:	2145
Score:	2015.00	Matches:	388
Percent Similarity:	80.87%	Conservative:	60
Best Local Similarity:	70.04%	Mismatches:	102
Query Match:	65.44%	Indels:	4
DB:	20	Gaps:	2

US-10-023-515-2 (1-581) x US-10-233-933A-3 (1-2145)

```

QY      31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnIleLysGln 50
Db      238 GGGCCAGCTGCTGATGACACAGTAGAGACACAGAGCTGGATGGATCCGGGAGAGACA 297
QY      |||
QY      51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
Db      298 ACCACTGATCTGGAGAGACCGCTGCTGAAACATGTTCTCTGGGATCCCTATCTGCA 357
QY      |||
QY      71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeu 90
Db      358 CCTCCTTAGGGGCCCTGCGATTTAAGCAACCAAGCCTGCTGCGCGGAGATACCTTC 417
QY      |||
QY      91 ArgGluAlaThrSerTrpProAsnLeuCyAspGlnAsnSerGluTrpLeuLeuAsp 110
Db      418 CGAATGCCACATCTTACCTTAATATATGCTTCCAGACCTTAAGATGGCTGCTCTAT 477
QY      |||
QY      111 GlnHisMetLeuLysValHisTrpProLysPheGlyValSerGluAspCyAspLeuTrpLeu 130
Db      478 CAACAGTTCTCAAGAGTGGTTACCCCAAAATTGMAAGCGTCGGAAGACTGCTGTACCTT 537
QY      |||
QY      131 AsnIleTrpAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhe 150
Db      538 AACATCTATGGCCGACCCCATGCGGACCAATGCTTCAACTCCCTGTATGATGTGTGTC 597
QY      |||
QY      151 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
Db      598 CCGGGGGTGCCTTCAAAATAGGCGCTCACTTCTCTGATGGTCCGGCTGTGCTGCC 657
QY      |||
QY      171 TyrGluAspValLeuValValValGlnTrpArgLeuGlyLysPheGlyPhePheThr 190
Db      658 TACGAGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
QY      |||
QY      191 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSer 210
Db      718 ACAGGGATGACAGTCCCGGGGGAATGGGCTTGTGAGACAGATGGCTGCCCTAACCC 777
QY      |||
QY      211 TrpValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230
Db      778 TGGTCCGGGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
QY      |||
QY      231 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 250
Db      838 GAGTCAGCGGGAGCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
QY      |||
QY      251 PheHisLysAlaIleMetGluSerGlyValAlaIleIleProTrpLeu---GluAlaHis 269
Db      898 TTCCACAAAGCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
QY      |||
QY      270 AspTrpGluLysSerGluAspLeuGlnValValAlaHisPheCyGlyAsnAsnAlaSer 289
Db      958 GGTATGAGAGAGAGAGAGATTTTGCAGTCTGCTGCGGCTATCTGTGTGTCATGCTCT 1017
QY      |||
QY      290 AapSerGluAlaLeuLeuArgCyAspLeuArgThrLysProSerLysGluLeuLeuThrLeu 309
Db      1018 GACTGTGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
QY      |||
QY      310 SerGlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluPro 329
Db      1078 AGCAAGAACTCAACGTTTTCATTCAGTGAATGATGATGATGATGATGATGATGATGATGAT 1137
QY      |||
QY      330 LeuAapLeuSeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAlaSerHis 349

```



```
Db 1138 GTAGCCCTATTGACTCAAAAAGCATTTAATTCAGTTCCTTATCATCGGAGTCAATAAC 1197
Qy 350 HIGGLUCYSGYRHELEULEUPROMETLYSGLUALAPROGUULEUSERGLYSERASN 369
Db 1198 CACGAGTGTGCTTCTCTG--TCCAGGAGTTTTCTAGATCCCTGGGCGCTCCAAC 1254
Qy 370 LYSERLEUALALEUHI1LEUHI1LEUHI1LEUHI1SERLEUPROGUINLYRLEUHI1S 389
Db 1255 AGGTCTGGGCTCTACTAGTACACACGCTCTGAATTTCCACCCAGTATTGGCAC 1314
Qy 390 LEUVAL1AANGUITYRPHEN1ASRPLYSHISERLEUTHRGUULEARSPSERLEU 409
Db 1315 CTGTGGCTAGTACTTACTTACACACAGCACTCCCGGTGAATAACAGATAGTTT 1374
Qy 410 LEUASRLEULEUGLYASRVAL1PHEPHEVAL1PROALALEU1LEUHI1AARGLYRHS 429
Db 1375 CTGGACTGCTTGGAGATGTCTCTTTGTGTGCTGGGTGTGACAGTCAATATCAAT 1434
Qy 430 ARGASRVAL1AGLYALAPROVAL1TYRPHETRYGLUPHEARGH1ASRPROGUINCYSPHEGU 449
Db 1435 AGAGATGTGTGTGACCTGTCTACTTCTATGAGTTTCAACACCCGCCAGTCTTAAC 1494
Qy 450 APTTHLYSRPROVAL1PHEVAL1LYSALASRHI1AASRGLUVAL1ARGPHEVAL1PHEGLY 469
Db 1495 GACACAGGCGCACTTCTGTGAAGCCGATCACTGTGATGAATCCGCTTCTTTTGA 1554
Qy 470 GLVAL1PHELEULEUGLYASR1LEVAL1METPHEGLUGLYALATHRGUUGUUGLYLEU 489
Db 1555 GGTCCTTCTCTGAAGCGACATTTGTCATGTTGMAAGAGCCAGAGAGAGAAATTG 1614
Qy 490 LEUSERARGLYMETMETLYSTYRTPRALATHRPHEN1AARGTHLYASRPROANGLY 509
Db 1615 CTGAGGAGGAAATGATGAGTACTGAGGCACTTGTGCGAGCCGGGAGACCTTAACGG 1674
Qy 510 AENASRLEUSERLEUTRPROVAL1TYRASNLEUTHRGUUGUUGLYLEUHI1LEUASRLEU 529
Db 1675 GAAGGTGTCTCTGTGGCGCACTACACCCAGAGCGAGCACTGAGCTGGAATTG 1734
Qy 530 AENMETSERLEUGLYGLNARGLEULEUGUUPROARGVAL1ASRPHETRTPH1SERTHILE 549
Db 1735 AGTGTAGCGGTGGACAGAACTGAAAGGAGGAGAGTTTGGATGAATCCATT 1794
Qy 550 --PROLEU1LEUSERALASERASRMETLEUHI1SERPROLEUSERLEUTHRPH1 569
Db 1795 GTCCCTGATACCCCCCACTTCAGAGGCCCTTCCCAAGTCTCTTNTCCCTTACTCTCC 1854
Qy 569 EUSERLEULEUGLNPROBPHEPHEPHECYVAL1PRO 581
Db 1855 TTCTTTGCTCCCGCTGCTTNTTTTTCGTGCTCA 1892

RESULT 14
US-10-233-933A-1
; Sequence 1, Application US/10233933A
; Publication No. US20040214171A1
; GENERAL INFORMATION:
; APPLICANT: Yamazaki, Tetsuo
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
; FILE REFERENCE: SHIG PRO2US006
; CURRENT APPLICATION NUMBER: US/10/233, 933A
; PRIOR FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: JP2002-057908
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Felis catus
US-10-233-933A-1
Alignment Scores:
Pred. No.: 6,5e-228 Length: 1629
```

```
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 64.47% Indels: 2
DB: 20 Gaps: 2

US-10-023-515-2 (1-581) x US-10-233-933A-1 (1-1629)
Qy 31 GLYPROSERALAGLUGLYPROGUINARGANTRARGLEUGLYTRP1LEGUINLYSGU 50
Db 64 GGGCCAGCGTGTGATGACACAGTAGAGAGCACAGGCTGGATGGGTGCGGGAGAGCA 123
Qy 51 VALTHRVALLLEUGLYSERPROVAL1PROVAL1ENVAL1PHELEUGLYVAL1PROPHEN1A 70
Db 124 ACCACTGTACTGGGAGAGCACCGCTGCTGTGAACATGTTCTCGGAGATCCCTATAGTCGA 183
Qy 71 PROPROLEUGLYSERLEUARGPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 90
Db 184 CCTCTCTAGGAGCCCTTGCATTTAAGCAACCAAGGCTGCTGCTGCGGAGATGACTTC 243
Qy 91 ARGGLUALATHRSERYTRPROASNLEUCYBLEUGLNANSERGUUTRPLEULEUASR 110
Db 244 CGAATGCGACATCTACCTTAATTAATGCTTCAGAGACTTAAGATGGCTGTCTCTAT 303
Qy 111 GLNH1SMETLEUYSVAL1H1STYRPROLYRPH1GLYVAL1SERGLUASRPLYRLEU 130
Db 304 CACACGTTCTCAAGAGCGTTTACCCCAATGGAGACCGTCCGAGACCTGCTGACTT 363
Qy 131 AEN1LETRYALAPROVAL1H1ASRTHRGY1SERLYSRLEUPROVAL1LEUVAL1TRPHE 150
Db 364 AACATCTATGGCGCAGCCCATGGGAGCAATGGCTCCCACTCTCTGTCAATGTGTGTC 423
Qy 151 PROGLYGLYALPHELYTRHGLYSERALASER1LEPHEASRGLYSERALALEU1A1A 170
Db 424 CCGGGGTGCTTCAAGATGGCTCAGCTTCTCTCATGATGGTCCGCTTGGCTGCTCC 483
Qy 171 TYRGLUASRVAL1LEUVAL1VAL1GLNTRYARGLEUGLY1LEPHEGLYRPHETNR 190
Db 484 TAGAGAGACGTGTGATTCGTGACTACCCAGTACCGGCTTGAATATTTGGTTTTTCAG 543
Qy 191 THTRPARGLNH1A1APROGUINATRPVAL1PHELYASRGLNVAL1A1ALEUSER 210
Db 544 ACAGGGATGAGCATCCCGGGGAACTGGGCTTCTGAGACAGGTGCTGCCCTTAAC 603
Qy 211 TRPVAL1GLNLYASN1LEGUPHEPHEGLYGLYASRPROSERVAL1THR1LEPHEGLY 230
Db 604 TGGGTCCGGAGCAACATCGAGTCTTGGTGTGACCCAGCTCCGTGACCATTTTGA 663
Qy 231 GLUSERAL1GLYAL1ESERVAL1SERSEULEU1LEUSERPROMERAL1ASRGLYU 250
Db 664 GAGTCAGCGGAGCCATCATGTTTCCAGCTCATTTCTCCCCCATAGCCAAATGGCTTA 723
Qy 251 PHEH1LYVAL11EMETGLUSERGLYVAL1A11E1LEPROTYRLEU--GLUALH1S 269
Db 724 TTCACAAAGCCATCATGAGAGTGGGGGCACTCCCTTACTAGATGAGACCCCT 783
Qy 270 ASRTRYGLYLSERGLUASRLEUGLNVAL1VAL1AH1SPHECYGLYASRASN1ASR 289
Db 784 GGTGATGAGAGAGAAAGATTTGCAAGTGTCTGGCGGTATGTGTGTTGGCATGCTCT 843
Qy 290 AASRERGLUAL1LEULEUARGCYLEUARGHLYRPH1PROSERLYSGU1LEUUTHR1 309
Db 844 GACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
Qy 310 SERGLNLYSRTHLYSERPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHEPHE 329
Db 904 AGCAAGAACTACGTTTCCATTTCCAGTGTATGATGATCTTTTCTTTCGTGATGAGCT 963
Qy 330 LEUASRLEUSERGLNLYSAL1PHELYVAL1LEPROSER1LE1GLYVAL1ASN1 349
Db 964 GTAGCCCTATTGACTCAAAAAGCATTTAATTCAGTCTCTTATCATCGGAGTCAATAAC 1023
Qy 350 HIGGLUCYSGYRHELEULEUPROMETLYSGLUALAPROGUULEUSERGLYSERASN 369
```

```
|||||
Db 1024 CACGAGTGTGCTTCTCTG---TCCACGGAGTTTCTGAGATCTCGGGGGCTCCAC 1080
Qy 370 LySerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProlGlnIleuHis 389
Db 1081 AGGCTCTGGGCGCTTCTACTAGTACACAGTTCCGAATATTTCCACCCAGATTTTGCAC 1140
Qy 390 LeuValAlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeu 409
Db 1141 CTGTGGCTGTGATCTACTTCTTACACAAAGCACTCCCGGTTGAATATCGAGATGTTT 1200
Qy 410 LeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHis 429
Db 1201 CTGAGCTGTCTGAGAGATGTGTCTTTGTGGTCCCTGGGGGTGTACAGCTCGATATCAT 1260
Qy 430 ArgAspAlaGlyAlaProValTyrPheTyrGluPheAsnGlnAspArgProGlnCysPheGlu 449
Db 1261 AGAGATGCTGGTGCACTGTCTACTTCTATGAGTTTCAACACCCCGCCAGTGTAAAC 1320
Qy 450 AspThrLysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGly 469
Db 1321 GACACGAGGCGCAAGCTTGTGTAAGCCGATCACCTGATGAATCCGCTTGTCTTTGGA 1380
Qy 470 GlyAlaPheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeu 489
Db 1381 GGTGCTTCTCGAAGGCGACATTTGTCAATGTTCCAGAGGCGACCGAGAGAGAAATTTG 1440
Qy 490 LeuSerArgLysMetMetLysTyrTrrAlaThrPheAlaArgThrGlyAsnProAsnGly 509
Db 1441 CTGAGCGAGGAAGATGATGAGTACTGGGCGCAACTTGTCTCGGACCGGGGAGCCCTTAACGGG 1500
Qy 510 AsnAspLeuSerLeuTrrProAlaTyrAsnLeuThrGluGlnIleGluLeuAspLeu 529
Db 1501 GAAGGTGTGCTCTGTGGCCAGCTTACACCCAGAGCGAGTACTTAAGCTGATTTTG 1560
Qy 530 AsnMetSerLeuGlyGlnArgLeuLysGluProArgValAspPheTrrThrSerThrIle 549
Db 1561 AGTGTGAGCGTGGGACAAATCTGAAGGAGCAAGAGTGGAGTTTGAATGAATCCATT 1620

RESULT 15
US-10-451-168-42
; Sequence 42, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1071
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-42

Alignment Scores:
Pred. No.: 1,63e-203 Length: 1071
Score: 1780.00 Matches: 340
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.81% Indels: 0
DB: 18 Gaps: 0

US-10-023-515-2 (1-581) x US-10-451-168-42 (1-1071)
Qy 242 IleLeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGluSerGlyValAla 261
Db 49 ATACTGTCTCCATGCGCCAAAGGCTTATTCACAAAGCATGAGAGAGGGGTGACC 108
Qy 262 IleIleProTyrLeuGluAlaHisAspTyrGluLysSerGluAspLeuGlnValAla 281
Db 109 ATCATCCCTTACCTGAGGCCCATATGATGAGAGGTGAGACCTGACAGGTGTGCA 168
Qy 282 HisPheCysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuValGlyLeuArgThrLys 301
Db 169 CATTCTGTGGTAACAATGCGCTCAGACTGTAGGCGCTGTGAGGTGCTGAGGACAAAA 228
Qy 302 ProSerLysGluLeuLeuThrLeuSerGlnIleThrLysSerPheThrArgValAlaAsp 321
Db 229 CCTCCAAAGAGCGTGAACCTTCAGCCAGCAAAACAAAGCTTTTCACTCGAGTGTGAT 288
Qy 322 GlyAlaPhePheProAsnGluProLeuAspLeuSerGlnLysAlaPheLysAlaIle 341
Db 289 GGTGCTTCTTCTTAAGAGCCCTTAGATCTATGTTCTCGAAGAGCATTTAAAGCATTT 348
Qy 342 ProSerIleIleGlyValAsnAsnHisGluCysGlyPheLeuLeuProMetLysGlyAla 361
Db 349 CTTTCATCATCGAGGTCAATTAACCAAGATGTGGCTTCTGCTGCTTATGAGAGGCT 408
Qy 362 ProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisIleGlnAsnIleLeu 381
Db 409 CTTGAGATCTTCAGTGGTCCAAACAGTCCCTTGCCCTCAATCGATACAAACATCTGTG 468
Qy 382 HisIleProProlGlnIleuHisLeuValAlaAsnGluTyrPheHisAspLysHisSer 401
Db 469 CACATCCGCGCTCAGATATTTGCACCTTGTGCTAATGAATCTTCCAGACAGACACTCC 528
Qy 402 LeuThrGluIleArgAspSerLeuLeuAspLeuGluAspValPhePheValAlaPro 421
Db 529 CTGACTGAATCCGAGACAGTCTTCTGGACTTGTGAGATGTGTCTTGTGTGCTCCCT 588
Qy 422 AlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTrrGluPhe 441
Db 589 GCACGTGACACAGCTCGATATCACAGATGCTGTGACCTGTCTACTTATGAGATT 648
Qy 442 ArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAspHisAla 461
Db 649 CGGACCGGCGCTCAGTGTGTAAGACAGAGCAGCTTTGTTCMAAGCCGACCACTCT 708
Qy 462 AspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGlu 481
Db 709 GATGAAGTCCGCTTGTGTGCTGTGCTTCTTCTGMAAGGGGACATTTGTTATGTTGAA 768
Qy 482 GlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysTyrTrrAlaThrPhe 501
Db 769 GGAGCGACGAGAGGAGAAAGTTACTGAGCGGAAAGATGAATATGAGGCTACCTTT 828
Qy 502 AlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrrProAlaTyrAsnLeuThr 521
Db 829 GCTGGAACGGGAATCTTAAGGAGACCACTGTCTGTGTGCGCAGCTTAATCTGACT 888
Qy 522 GluGlnIleLeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLysGluProArg 541
Db 889 GAGCAGTACTTCCAGCTGAGCTTGAACATGAGCTTGGACAGAGACTCAAGAAACCCGG 948
```

Qy	542	VaIAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAapMetLeuHisSer	561
Db	949	GTGGATTTTGGACACGACCAATCCCTGATCTGTGCTCCGACATGCTCCACAGT	1008
Qy	562	ProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAlaPro	581
Db	1009	CCTCTTTCCTTCTTAACTTCTCTCTCTCCAGCCTTCTTTCTTTGTGCTCCT	1068

Search completed: June 15, 2005, 15:44:59
Job time : 994 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 08:55:45 ; Search time 87 Seconds
(without alignments)
2582.848 Million cell updates/sec

Title: US-10-023-515-2
3079

Perfect score: 1 MPGLTSSASQWCFLLIQP.....PLSSLTFLSLQPPFFCAP 581
Sequence:

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq.16Dec04:*

1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3079	100.0	581 5 ABB79537	Abb79537 Human car
2	3079	100.0	581 8 ABO89094	Abd89094 Human uro
3	2915	94.7	575 5 ABP61004	Abp61004 Novel hum
4	2900	94.2	642 5 AAE25025	AAe25025 Human dnu
5	2900	94.2	642 8 ABM84114	ABm84114 Human dia
6	2607.5	84.7	581 5 ABP61005	Abp61005 Novel hum
7	2606	84.6	525 6 ADA54807	Ada54807 Human pro
8	2597.5	84.4	581 6 ABUS54539	Abus54539 Human NOV
9	2586.5	84.1	618 5 ADR19663	Adr19663 Human dnu
10	2454	79.7	469 7 ADB64065	Adb64065 Human pro
11	1985	64.5	542 8 ADF50145	Adf50145 Cat cauxi
12	1985	64.5	542 8 ADF50147	Adf50147 Cat cauxi
13	1780	57.8	356 5 ABP61006	Abp61006 Novel hum
14	1262.5	41.0	561 5 AAE20911	AAe20911 Rat carbo
15	1230	39.9	549 3 AAB58981	AAb58981 Breast an
16	1230	39.9	550 8 AOG98219	Aog98219 Human int
17	1230	39.9	559 4 AAB31700	AAb31700 Protein e
18	1230	39.9	559 4 ABUS3223	Abus3223 Human met
19	1230	39.9	559 6 ABB82851	ABb82851 CES2 rela
20	1230	39.9	559 6 AAE33960	AAe33960 Human car
21	1230	39.9	559 7 ADD46521	Ad446521 Human pro
22	1221	39.7	306 5 AAE20909	AAe20909 Human car
23	1213.5	39.4	561 7 ADD46519	Ad446519 Rat Prote
24	1191	38.7	607 4 ABUS3222	Abus3222 Human met
25	1186.5	38.5	554 8 ADI79890	Adi79890 Mouse liv

26	1181	38.4	583 4 ABG10273	ABg10273 Novel hum
27	1179.5	38.3	571 3 AAY71107	Aay71107 Human Hyd
28	1179.5	38.3	571 4 AAU12442	AAu12442 Human PRO
29	1179.5	38.3	571 4 AAE04101	AAe04101 Human Gen
30	1179.5	38.3	571 5 AAU83696	AAu83696 Human PRO
31	1179.5	38.3	571 5 ABB84949	ABb84949 Human PRO
32	1179.5	38.3	571 5 ABG64341	ABg64341 Human alb
33	1179.5	38.3	571 6 ABO17886	ABo17886 Novel hum
34	1179.5	38.3	571 6 ABU69108	ABu69108 Human PRO
35	1179.5	38.3	571 6 ABU80843	ABu80843 Human PRO
36	1179.5	38.3	571 6 ABO33809	ABo33809 Novel hum
37	1179.5	38.3	571 6 ABUB1140	ABuB1140 Human PRO
38	1179.5	38.3	571 6 ABO19424	ABo19424 Human sec
39	1179.5	38.3	571 6 ABU66840	ABu66840 Human PRO
40	1179.5	38.3	571 6 ABUS9921	ABus9921 Novel sec
41	1179.5	38.3	571 6 ABU69085	ABu69085 Human PRO
42	1179.5	38.3	571 6 ABO25111	ABo25111 Human sec
43	1179.5	38.3	571 6 ABU82152	ABu82152 Novel hum
44	1179.5	38.3	571 6 ABU67116	ABu67116 Human sec
45	1179.5	38.3	571 6 ABU81549	ABu81549 Human sec

ALIGNMENTS

RESULT 1
ABB79537

ID ABB79537 standard; protein; 581 AA.

AC ABB79537;

DT 23-SEP-2002 (first entry)

DB Human carboxylesterase family member 53010.

KW Carboxylesterase; enzyme; human; analgesic; nootropic; antiinflammatory; diagnosis; therapy.

KX Homo sapiens.

OS Homo sapiens.

FX Key

FT Peptide

FT Protein

FT Domain

FT Region

FT Active-site

FT WO200250256-A2.

PD 27-JUN-2002.

PF 18-DEC-2001; 2001WO-US049075.

PR 18-DEC-2000; 2000US-0256369P.

PR 28-MAR-2001; 2001US-0279508P.

XX (MILL-) MILLENNIUM PHARM INC.

PA Curtis Rad, Silos-Santiago I;

XX WPI; 2002-547936/58.

DR N-PSDB; ABN84302.

XX 53010 nucleic acids, useful for diagnosing and treating e.g. vascular diseases, autoimmune diseases, or neurodegenerative diseases; as surrogate markers, in tissue typing and chromosome mapping.


```

Db 121 FGVEDCLYINLYAPAHADYGSKLPLVWVFGCAFTGSASIFDGSALAAYEDVLVVVQ 180
QY YRLGIFGPTTDDOAPGNWAFKDOVALSWQKNIIEFGGPPSSVTTFGESAGAISSVS 240
Db 181 YRLGIFGPTTDDOAPGNWAFKDOVALSWQKNIIEFGGPPSSVTTFGESAGAISSVS 240
QY 241 LILSPMAKGLFPAKAIMESGVAIIPYLEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRT 300
Db 241 LILSPMAKGLFPAKAIMESGVAIIPYLEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRT 300
QY 301 KPSKELLTLISQKTKSPTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKE 360
Db 301 KPSKELLTLISQKTKSPTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKE 360
QY 361 APEILSGNSKSLALHLIQLNIHLIIPQYVHLVANEYFHDGSLTEIRDSLDDLGDVFPV 420
Db 361 APEILSGNSKSLALHLIQLNIHLIIPQYVHLVANEYFHDGSLTEIRDSLDDLGDVFPV 420
QY 421 PALITARYHRDAGAVYFYEFPNRPQCEEDTKPAFVKAHDADVAFVFGAFLKGDIVWF 480
Db 421 PALITARYHRDAGAVYFYEFPNRPQCEEDTKPAFVKAHDADVAFVFGAFLKGDIVWF 480
QY 481 EGATEEEKLSHKMKMYATFARTGNPNGNDSLMPAYNLTEQYLQDLNMSLGRLKEP 540
Db 481 EGATEEEKLSHKMKMYATFARTGNPNGNDSLMPAYNLTEQYLQDLNMSLGRLKEP 540
QY 541 RVDFTWSTIPLILASDMLHSPSLTLPLSLQPPFFCAP 581
Db 541 RVDFTWSTIPLILASDMLHSPSLTLPLSLQPPFFCAP 581

```

RESULT 3
ABP61004
ID ABP61004 standard; protein; 575 AA.

XX AC ABP61004;
XX DT 10-SEP-2002 (first entry)
XX DE Novel human protein. SEQ ID 91.

XX Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian;
XX nootropic; neuroprotective; immunosuppressive; haemostatic;
XX antiinflammatory; cardiant; antilicer; vaccine; cancer; infection;
XX cerebroprotective; anorectic; metabolic; metabolic; cancer; infection;
XX wound healing disorders; atherosclerosis; Parkinson's disease;
XX Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
XX inflammation; neoplastic disease; nervous system disorder;
XX cardiovascular disorders; pancreatitis; respiratory disorder;
XX hyperproliferation; systemic autoimmune disease; hyper-immunity;
XX developmental abnormality; gastrointestinal ulceration; neuropathy;
XX haematological disease; metabolic disease; sperm dysfunction;
XX thyroid disorder; hypothyroidism; brain damage; colitis;
XX cone photo-transduction deficiency; neurological disease; stroke;
XX angiodysplasia; ovulation disorder; spinal cord; thyroid gland; heart;
XX trachea; thymus; lymph node; muscular system; obesity; anorexia;
XX growth abnormality; precocious puberty.

XX Homo sapiens.
XX WO20025105-A1.

XX 27-JUN-2002.

XX 17-DEC-2001; 2001WO-US049232.

XX 19-DEC-2000; 2000US-0256710P.

XX 20-DEC-2000; 2000US-0257048P.

XX 09-JAN-2001; 2001US-0260482P.

XX 30-JAN-2001; 2001US-0264922P.

XX 06-FEB-2001; 2001US-0266797P.

XX 19-MAR-2001; 2001US-0276988P.

XX 04-APR-2001; 2001US-0281535P.

```

PR 08-MAY-2001; 2001US-0289622P.
XX (SMK ) SMITHLINE BECHAM CORP.
PA (SMK ) SMITHLINE BECHAM PLC.
PA (GLAX ) GLAXO GROUP LTD.
XX Agawwal P, Birkeland M, Cogswell JF, Kabnick KP, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX MPI: 2002-508784/54.
DR N-PSDB; AB086169.
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
XX Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX Claim 1(a); Page 312-313; 335pp; English.
XX PS The invention relates to an isolated polypeptide with signal sequences
XX CC which allow it to be secreted extracellularly or membrane associated. The
XX CC activity of polypeptides of the invention may be described as,
XX CC cytostatic, vulnary, antiarteriosclerotic, antiparkinsonian, nootropic,
XX CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
XX CC cardiant, antilicer, vaccine, antihypertensive, cerebroprotective, anorectic,
XX CC and metabolic. Polypeptides and polynucleotides of the invention are
XX CC useful in the treatment, or as a vaccine in the prevention of, cancer,
XX CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
XX CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
XX CC inflammation, neoplastic diseases, nervous system related disorders and
XX CC cardiovascular disorders, pancreatitis, respiratory disorder,
XX CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
XX CC developmental abnormality, gastrointestinal ulceration, neuropathy,
XX CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
XX CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
XX CC transduction deficiency, neurological diseases, stroke, angiodysplasia,
XX CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
XX CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
XX CC growth abnormalities, and alleviation of precocious puberty. The
XX CC sequences given in records ABP60965-ABP61019 represent novel human
XX CC proteins of the invention
XX SQ Sequence 575 AA;

```

Query Match 94.7%; Score 2915; DB 5; Length 575;
Beef Local Similarity 97.0%; Pred. No. 2.3e-263;
Matches 553; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

```

QY 12 WCFPLIOLPLIGHRQWGTGSAEGPQNTRLGWTQKQVTVLGSPPVNVFLGVPPAP 71
Db 14 MAIWTLAP-----TKGSABGPQNTRLGWTQKQVTVLGSPPVNVFLGVPPAP 65
QY 72 PLGSIIRTPNPASFWNDLREATSYPNICLQNSEWLLIDQMLKXHYRFGVSEDCLYLN 131
Db 66 PLGSIIRTPNPASFWNDLREATSYPNICLQNSEWLLIDQMLKXHYRFGVSEDCLYLN 125
QY 132 IYAPAHADYGSKLPLVWVFGCAFTGSASIFDGSALAAYEDVLVVVQYRLGIFGPTT 191
Db 126 IYAPAHADYGSKLPLVWVFGCAFTGSASIFDGSALAAYEDVLVVVQYRLGIFGPTT 185
QY 192 WDOHAPGNMAFADQVVALSWQKNIIEFGGPPSSVTTFGESAGAISSVSLLSPMAKGLF 251
Db 186 WDOHAPGNMAFADQVVALSWQKNIIEFGGPPSSVTTFGESAGAISSVSLLSPMAKGLF 245
QY 246 HKAIMESGVAIIPYLEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRTKPSKELLTLISQ 305
Db 252 HKAIMESGVAIIPYLEAHDEYKSEDLQVVAHFCGNNASDSBALRCLRTKPSKELLTLISQ 311
QY 312 KTKSFTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKEAPBIIISGNSKS 371
Db 306 KTKSFTRVVDGAFPPNEBLLISQAKFAIPSIIGVNNHECGFLLPMKEAPBIIISGNSKS 365
QY 372 LALHLIQLNIHLIIPQYVHLVANEYFHDGSLTEIRDSLDDLGDVFPVVALITARYARD 431

```


Db 572 DLFLPAPVNLTEQYQLDNLNLSLQRLKEPRVFTSTPILLSASDMLSPSSLTFLS 631

QY 571 LLQPPFFCAP 581

Db 632 LLQPPFFCAP 642

RESULT 5

ABM84114 standard; protein; 642 AA.

ABM84114;

18-NOV-2004 (first entry)

Human diagnostic and therapeutic pprotein SEQ ID NO:4363.

gene therapy; human diagnostic and therapeutic polynucleotide; dthp.

Homo sapiens.

MO2004023973-A2.

25-MAR-2004.

12-SEP-2003; 2003MO-US028227.

12-SEP-2002; 2002US-0410259P.

12-SEP-2002; 2002US-0410260P.

(INCY-) INCYTE CORP.

Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;

Hathshorne TA, Suchotolski MT, Altus CM, Plets SJ, Elder LV;

Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;

Stevens KM, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;

Peralta CH, Anderson SB, Rioux P, Shen EJ, Mu MC, Stuve LL;

Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirtson ES;

Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;

Pacutry S, Shi X, Suarez CJ;

MP1; 2004-329368/30.

N-PSDB; ACN42766.

New diagnostic and therapeutic polynucleotides and polypeptides, useful

in diagnosing a condition, disease or disorder associated with human

molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or

in gene mapping.

Claim 27; Page: 190pp; English.

The invention relates to novel diagnostic and therapeutic polynucleotides

selected from one of the 2722 sequences defined in the specification. A

polynucleotide of the invention may have a use in gene therapy. The human

diagnostic and therapeutic polynucleotides (dthp) or polypeptides may be

used to diagnose a particular condition, disease or disorder associated

with human molecules, e.g. cell proliferative disorders,

autoimmune/inflammatory disorder, developmental disorders, endocrine

disorder, neurological disorders, gastrointestinal disorders, or

infections caused by virus, bacteria, fungi or parasite. The dthp

molecules may also be used in genetic mapping, in identifying individuals

from minute biological samples, in detecting single nucleotide

polymorphisms, as molecular weight markers, and for somatic or germline

gene therapy. The present sequence represents a dthp protein of the

invention. Note: The sequence data for this patent is not represented in

the printed specification, but was obtained in electronic format directly

from WIPO at www.wipo.int/pct/en/sequences/listing.htm

Sequence 642 AA;

Query Match 94.2%; Score 2900; DB 8; Length 642;

Best Local Similarity 99.6%; Pred. No. 7e-262;

Matches 549; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 31 GPSAGPQNRNLIGTIOGKQVTVGSPPVNVVFLGVPAPPLGSLRFTNPQASPMPNL 90

Db 92 GPSAGPQNRNLIGTIOGKQVTVGSPPVNVVFLGVPAPPLGSLRFTNPQASPMPNL 151

QY 91 REATSYPLCLQNSWMLLDQMLKVHYPKGVSEDCYLNITYAPAHADTGSKLPLVWF 150

Db 152 REATSYPLCLQNSWMLLDQMLKVHYPKGVSEDCYLNITYAPAHADTGSKLPLVWF 211

QY 151 PGAFKTSASIFDGSALAAVBDVLYVVQYRLGIFGFTTWDQAPGNMAFKQVAAALS 210

Db 212 PGAFKTSASIFDGSALAAVBDVLYVVQYRLGIFGFTTWDQAPGNMAFKQVAAALS 271

QY 211 WQKNIEFFGDPSSVTIRGSGAGISVSLISPMAGLPHKAIMESGVALIPLYEAND 270

Db 272 WQKNIEFFGDPSSVTIRGSGAGISVSLISPMAGLPHKAIMESGVALIPLYEAND 331

QY 271 YKSEEDLVVAHFCGNNADEBALRCLRTKPSKELLTSQKTSFTRVVDGAFPPNEPL 330

Db 332 YKSEEDLVVAHFCGNNADEBALRCLRTKPSKELLTSQKTSFTRVVDGAFPPNEPL 391

QY 331 DLSQKAFKALPSITGVNNHCGFLPMKEAPELISGNSKSLAHLIONLIHIPQYHL 390

Db 392 DLSQKAFKALPSITGVNNHCGFLPMKEAPELISGNSKSLAHLIONLIHIPQYHL 451

QY 391 VANEYFHDKSLTETRDSDLDDGVFVVPALITARYHRAAGAPVYEFEPHRCQFED 450

Db 452 VANEYFHDKSLTETRDSDLDDGVFVVPALITARYHRAAGAPVYEFEPHRCQFED 511

QY 451 TKPAFVKADHAEVAFVFGAFFLKGDIVMEGATEEELSLRKKMKWATFARTGNPGN 510

Db 512 TKPAFVKADHAEVAFVFGAFFLKGDIVMEGATEEELSLRKKMKWATFARTGNPGN 571

QY 511 DLSTMPAVNLTEQYQLDNLNLSLQRLKEPRVFTSTPILLSASDMLSPSSLTFLS 570

Db 572 DLSTMPAVNLTEQYQLDNLNLSLQRLKEPRVFTSTPILLSASDMLSPSSLTFLS 631

QY 571 LLQPPFFCAP 581

Db 632 LLQPPFFCAP 642

RESULT 6

ABP61005 standard; protein; 581 AA.

ABP61005;

10-SEP-2002 (first entry)

Novel human protein. SEQ ID 92.

Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;

nootropic; neuroprotective; immunosuppressive; haemostatic;

antiinflammatory; cardiant; antilicer; vincide; antithyroid;

cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;

wound healing disorder; atherosclerosis; Parkinson's disease;

Alzheimer's disease; autoimmune disorder; haematopoietic disorder;

inflammation; neoplastic disease; nervous system disorder;

cardiovascular disorders; pancreatitis; respiratory disorder;

hyperlipidation; systemic autoimmune disease; hyper-immunity;

developmental abnormality; gastrointestinal ulceration; neuropathy;

haematological disease; metabolic disease; sperm dysfunction;

thyroid disorder; hypothyroidism; brain damage; colitis;

cone photo- transduction deficiency; neurological disease; stroke;

angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;

trachea; thymus; lymph node; muscular system; obesity; anorexia;

growth abnormality; precocious puberty.

Homo sapiens.

WO20025105-A1.

XX 27-JUN-2002.
 PD 17-DEC-2001; 2001WO-US049232.
 XX
 PF 19-DEC-2000; 2000US-0256710P.
 XX 20-DEC-2000; 2000US-0257048P.
 PR 09-JAN-2001; 2001US-0260482P.
 PR 30-JAN-2001; 2001US-0264922P.
 PR 06-FEB-2001; 2001US-0266787P.
 PR 19-MAR-2001; 2001US-0276988P.
 PR 04-APR-2001; 2001US-0281535P.
 PR 08-MAY-2001; 2001US-0289622P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PA (GLAX) GLAXO GROUP LTD.
 XX
 PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,
 PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q,
 XX N-PSDB; ABQ86170.
 DR WPI; 2002-508784/54.
 XX
 PT Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wound, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
 XX
 PS Claim 1(a); Page 313-314, 335pp; English.
 PS
 XX The invention relates to an isolated polypeptide with signal sequences
 CC which allow it to be secreted extracellularly or membrane associated. The
 CC activity of polypeptides of the invention may be described as,
 CC cytostatic, vulnerary, antihypertensive, antiparkinsonian, nootropic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 CC cardiact, anticancer, antiviral, antihypertensive, cerebroprotective, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder,
 CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
 CC haematological abnormalities, gastrointestinal ulceration, neuropathy,
 CC developmental abnormalities, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, angiogenesis,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABP60965-ABP61019 represent novel human
 CC proteins of the invention
 CC
 XX Sequence 581 AA;
 XX
 SQ
 Query Match 84.7%; Score 2607.5; DB 5; Length 581;
 Best Local Similarity 91.4%; Pred. No. 1.4e-234;
 Matches 499; Conservative 7; Mismatches 25; Indels 15; Gaps 3;
 QY 12 WCFFLIQIPLIGHQKQKTPSAEGPQPNRLGNIQKQVTVLGSPVNVVFLGVPPAP 71
 DB 14 WAIWVLAAP-----TKGPSAEGPQPNRLGNIQKQVTVLGSPVNVVFLGVPPAP 65
 QY 72 PLGSIKRTNQPASPMQNLREATGYPMLCQNSWILLDDHMLKMYPRKGVSEDCLYLN 131
 DB 66 PLGSIKRTNQPASPMQNLREATGYPMLCQNSWILLDDHMLKMYPRKGVSEDCLYLN 125
 QY 132 IYAPAHADTGSKLPVLVWFFPGAFKXTGSASIFDGSALAAEDVLVWVQYRLGIFGFTT 191
 DB 126 IYAPAHADTGSKLPVLVWFFPGAFKXTGSASIFDGSALAAEDVLVWVQYRLGIFGFTT 185
 QY 192 WDOHAGNNAFKQVVALSVVQKNIEFGQDPSSVTIFGESAGATSVSSILISPMAKGLF 251

DB 186 WDOHAGNNAFKQVVALSVVQKNIEFGQDPSSVTIFGESAGATSVSSILISPMAKGLF 245
 QY 252 HKAIMESGVALIPYLEADHYEKSEBDLOVVAHFCGNNADESEALRLCLTKPSKEILLTSQ 311
 DB 246 HKAIMESGVALIPYLEADHYEKSEBDLOVVAHFCGNNADESEALRLCLTKPSKEILLTSQ 305
 QY 312 KTKSFTRVVDGAFPFNEBPLDLSQKAFAPISIIIGVNNHECGFLLPMKE--APEILSGSN 369
 DB 306 KTKSFTRVVDGAFPFNEBPLDLSQKAFAPISIIIGVNNHECGFLLPMVRIIAVTAFTSN 365
 QY 370 KSLAL-----HLIQNLIIPQOYLHVANEFHDKHSITETIRSDLLDLGVFFVVALI 424
 DB 366 KDAALASTAGHFRHGHQIIPQOYLHVANEFHDKHSITETIRSDLLDLGVFFVVALI 425
 QY 425 TARYHRDAGAPVYFYEPRHRPQCFEDTKPAFVKADHDEVRFVVGAFPLGDTIMEFGAT 484
 DB 426 TARYHRDAGAPVYFYEPRHRPQCFEDTKPAFVKADHDEVRFVVGAFPLGDTIMEFGAT 485
 QY 485 EBEKILSRKMKKWTATPARGNPNNDLSLMPAYNLTEOYLQDLNNSLQRLKEPRVDF 544
 DB 486 EBEKILSRKMKKWTATPARGNPNNDLSLMPAYNLTEOYLQDLNNSLQRLKEPRRDV 545
 QY 545 WTSTRIP 550
 DB 546 WVTGYP 551

RESULT 7
 ID ADA54807 standard; protein; 525 AA.
 XX ADA54807;
 AC
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human protein, SEQ ID 2375.
 XX
 KM Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KM Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KM inflammatory disease; osteoporosis; neurological disease.
 OS Homo sapiens.
 XX
 PN EPI293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-0006586.
 PR 14-SEP-2001; 2001UP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) RES ASSOC BIOTECHNOLOGY.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehi S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maunoh Y;
 XX
 DR WPI; 2003-395539/38.
 DR N-PSDB; ADA53168.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 PS
 XX Claim 14; SEQ ID NO 2375; 205pp; English.
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 525 AA;
 SQ Query Match 84.6%; Score 2606; DB 6; Length 525;
 Best Local Similarity 90.9%; Pred. No. 1.7e-234;
 Matches 501; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

QY 31 GPSAEGPORNRLGWIQGVTVLGSPPVNVVFLGVPAPAPPLGSLRFTNPQAPAPMNL 90
 DB 25 GPSAEGPORNRLGWIQGVTVLGSPPVNVVFLGVPAPAPPLGSLRFTNPQAPAPMNL 84
 QY 91 REATSYPLCLQNSEWMLLDQMLKVHYPKFVSDCLYINITYAPAHADTGSKLEPVLPV 150
 DB 85 REATSYPLCLQNSEWMLLDQMLKVHYPKFVSDCLYINITYAPAHADTGSKLEPVLPV 144
 QY 151 PGAPKTKGSASIPDGSALAAVEDVTVVVOYRLGIFGFPTTWDQAPGNWAPKQVVALS 210
 DB 145 PGAPKTKGSASIPDGSALAAVEDVTVVVOYRLGIFGFPTTWDQAPGNWAPKQVVALS 204
 QY 211 WYQKNIEPFGDPPSSVTIFGSAGASISVSLISPMAGLPHKAIMESGVALIPLYEAMD 270
 DB 205 WYQKNIEPFGDPPSSVTIFGSAGASISVSLISPMAGLPHKAIMESGVALIPLYEAMD 264
 QY 271 YEKSEDLQVNAHFCGNNSDBALRLCRLTKRPSKELTLISQTKYSPTRVVDGAPFNEPL 330
 DB 265 YEKSEDLQVNAHFCGNNSDBALRLCRLTKRPSKELTLISQTKYSPTRVVDGAPFNEPL 324
 QY 331 DLISQKAPKAIPIIIGVNNHCGFLPMKEAPELISGNSKSLAHLIONIILHIPQYVHL 390
 DB 325 DLISQKAPKAIPIIIGVNNHCGFLPMKEAPELISGNSKSLAHLIONIILHIPQYVHL 384
 QY 391 VANEFPHDKSLTEIRDSLLDLGPFVFPVLPALITARVHRPAGAVFYEFRRHRQCPED 450
 DB 385 VANEFPHDKSLTEIRDSLLDLGPFVFPVLPALITARVHRPAGAVFYEFRRHRQCPED 424
 QY 451 TKPAPVNAHDAEVRFPVGAFLKGDIVMFGATEEKKLRSKMKKYATFARTGNPNGN 510
 DB 425 -----EGATEEKKLRSKMKKYATFARTGNPNGN 454
 QY 511 DLSTMPAVNLTQOYLQDLNNSLQRLKEPRVDFTSTIPILASDMLHSGLSLTLS 570
 DB 455 DLSTMPAVNLTQOYLQDLNNSLQRLKEPRVDFTSTIPILASDMLHSGLSLTLS 514
 QY 571 LLOPPEFFCAP 581
 DB 515 LLOPPEFFCAP 525

RESULT 8
 ABUS4639 standard; procein, 581 AA.
 ID ABUS4639;
 AC ABUS4639;
 XX
 DT 03-JUN-2003 (first entry)
 XX
 DB Human NOX polypeptide #98.
 XX
 XX Human; NOX; metabolic disorder; cardiomyopathy; diabetes; ASD;
 KW hypertension; congenital heart defect; aortic stenosis; valve disease;
 KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
 KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
 KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
 KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
 KW Parkinson's disease; immune disorder; haematopoietic disorder;
 KW haemophilia; hypercoagulation; Crohn's disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN MO200281498-A2.
 XX
 PD 17-OCT-2002.
 XX

PF 03-APR-2002; 2002WO-US010780.
 XX
 PR 03-APR-2001; 2001US-0281086P.
 PR 03-APR-2001; 2001US-0281136P.
 PR 05-APR-2001; 2001US-0281863P.
 PR 05-APR-2001; 2001US-0281906P.
 PR 06-APR-2001; 2001US-0282020P.
 PR 10-APR-2001; 2001US-0282930P.
 PR 10-APR-2001; 2001US-0282934P.
 PR 12-APR-2001; 2001US-0283512P.
 PR 13-APR-2001; 2001US-0283710P.
 PR 17-APR-2001; 2001US-0284234P.
 PR 19-APR-2001; 2001US-0285325P.
 PR 20-APR-2001; 2001US-0285381P.
 PR 20-APR-2001; 2001US-0285609P.
 PR 23-APR-2001; 2001US-0285748P.
 PR 23-APR-2001; 2001US-0285890P.
 PR 24-APR-2001; 2001US-0286068P.
 PR 25-APR-2001; 2001US-0286292P.
 PR 27-APR-2001; 2001US-0287213P.
 PR 02-MAY-2001; 2001US-0288257P.
 PR 29-MAY-2001; 2001US-0294164P.
 PR 30-MAY-2001; 2001US-0294484P.
 PR 18-JUN-2001; 2001US-0298952P.
 PR 19-JUN-2001; 2001US-0299237P.
 PR 19-JUN-2001; 2001US-0299276P.
 PR 12-SEP-2001; 2001US-0318750P.
 PR 25-SEP-2001; 2001US-0324800P.
 PR 25-SEP-2001; 2001US-0324802P.
 PR 27-SEP-2001; 2001US-0325684P.
 PR 17-OCT-2001; 2001US-0330143P.
 PR 14-NOV-2001; 2001US-0332131P.
 PR 14-NOV-2001; 2001US-0332240P.
 PR 14-NOV-2001; 2001US-0332779P.
 PR 21-NOV-2001; 2001US-0332115P.
 PR 04-DEC-2001; 2001US-0337621P.
 PR 03-JAN-2002; 2002US-0345783P.
 PR 16-JAN-2002; 2002US-0350251P.
 PR 02-APR-2002; 2002US-00114270.
 XX
 PA (CDRA-) CPRAGEN CORP.
 XX
 PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
 PI Paturajan M, Liu X, Gusev VY, Li L, Vernet CM, Zeehuseen BD;
 PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
 PI Padigaru M, Shinkets RA, Gangoli EA, Taupier RJ, Caeman SJ, Ji W,
 PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
 PI MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
 PI Ellerman K;
 XX
 DR WPI, 2003-046858/04.
 DR N-PsDB; ABX72267.
 XX
 PT New isolated NOX polypeptide useful for treating atherosclerosis,
 PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
 PT neurodegenerative disorders, Alzheimer's disease and cancer.
 XX
 PS Claim 1; Page 303; 666pp; English.
 XX
 XX The invention relates to human polypeptides, termed NOX, and the
 CC polynucleotides encoding them. The polypeptides and polynucleotides are
 CC useful for diagnosing disease, and screening for potential therapeutic
 CC agents. The sequences are useful for treating metabolic disorders,
 CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
 CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
 CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
 CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
 CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
 CC and cancer. Sequences ABUS4542-ABUS4647 represent human NOX polypeptides
 CC of the invention
 XX

SQ Sequence 581 AA;
 Query Match 84.4%; Score 2597.5; DB 6; Length 581;
 Best Local Similarity 94.3%; Pred. No. 1.2e-233;
 Matches 497; Conservative 4; Mismatches 17; Indels 9; Gaps 3;

QY 31 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPPAPPLGSRFTNPOASPMWDL 90
 DB 27 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPPAPPLGSRFTNPOASPMWDL 86
 QY 91 REATSYPNLCLQNSEWLLDQHLKHYPRKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 150
 DB 87 REATSYPNLCLQNSEWLLDQHLKHYPRKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 146
 QY 151 PGGAFTGSAISIFDGSALAAEDVYVWVVOYRLGIGFTTWDQAPGMAFKDOYVALS 210
 DB 147 PGGAFTGSAISIFDGSALAAEDVYVWVVOYRLGIGFTTWDQAPGMAFKDOYVALS 206
 QY 211 WVOKNIEFGDPSSVTIFGESAGAISVSLLISPMAGLPHKAIMESGVATIPYLEAH 270
 DB 207 WVOKNIEFGDPSSVTIFGESAGAISVSLLISPMAGLPHKAIMESGVATIPYLEAH 266
 QY 271 YKSESDLOVVAHFCGNMADSEALRLCLRTKPSKELTLTISQTKSFTTVVDAFPNEBL 330
 DB 267 YKSESDLOVVAHFCGNMADSEALRLCLRTKPSKELTLTISQTKSFTTVVDAFPNEBL 326
 QY 331 DLSQKAFKAIPIISIGVNNHECGFLPMKE--APEILSGSNKSLAL-----HLQNIHLI 383
 DB 327 DLSQKAFKAIPIISIGVNNHECGFLPMKRIAAHTATPNSNDALASTAGHFHRR--HI 384
 QY 384 PPOYLHLVANEFYHDKHSLTEIRDSLLDLGDFVFPVPLITARYHRDAGAVYEFERH 443
 DB 385 PPOYLHLVANEFYHDKHSLTEIRDSLLDLGDFVFPVPLITARYHRDAGAVYEFERH 444
 QY 444 RPOCFEDTPKAFVKADHDEVAFVFGAFLKGDIVMFGATBEELLSRKMMKWATPAR 503
 DB 445 RPOCFEDTPKAFVKADHDEVAFVFGAFLKGDIVMFGATBEELLSRKMMKWATPAR 504
 QY 504 TGNPNGNDLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDFTSTIP 550
 DB 505 TGNPNGNDLSLMPAYNLTEOYLQDLNMSLQRLKEPRVDFTSTIP 551

RESULT 9
 ADR19663
 ID ADR19663 standard; protein; 618 AA.
 AC ADR19663;
 DT 07-OCT-2004 (first entry)
 XX Human drug metabolizing enzyme (DME)-2 protein sequence.
 DB Human drug metabolizing enzyme (DME)-2 protein sequence.
 XX drug metabolizing enzyme; DME; cytochrome; immunosuppressive;
 KW antiinflammatory; endocrine; ophthalmologic; gastrointestinal;
 KW hepatocellular cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; endocrine disorder; eye disorder;
 KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;
 KW human.
 XX Homo sapiens.
 OS
 PN WO200226988-A2.
 PD 04-APR-2002.
 XX 28-SEP-2001; 2001MO-US030662.
 PF 29-SEP-2000; 2000US-0236947P.
 PR 06-OCT-2000; 2000US-0238864P.
 PR 20-OCT-2000; 2000US-0242333P.
 PR 09-NOV-2000; 2000US-0242581P.
 PR 16-NOV-2000; 2000US-0249519P.

PR 22-NOV-2000; 2000US-0252834P.
 PR 30-NOV-2000; 2000US-0250567P.
 PA (INCY-) INCYTE GENOMICS INC.
 XX Azinza Y, Baughin MR, Borowsky ML, Ding L, Dugan BW,
 PI Biliote VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan FA;
 PI Lai P, Lee BA, Lu DAM, Nguyen DB, Arvitz C, Policky JL, Ramkumar J;
 PI Ring HZ, Sanjanwala MS, Tang YF, Tribouley CM, Narinder WK,
 PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
 DR MPI: 2002-362498/39.
 DR N-ESDB; ADR19681.
 XX Human drug metabolizing enzymes, useful in the diagnosis and treatment of
 PT disorders associated with aberrant (DME) activity, e.g., cancer and
 PT autoimmune disorders.
 XX Claim 1; SEQ ID NO 2; 142bp; English.
 XX This invention relates to novel drug metabolizing enzymes (DME) and the
 CC nucleotide sequences which encode them. The invention may be useful for
 CC the development of compounds with a cytochrome, immunosuppressive,
 CC antiinflammatory, endocrine, ophthalmologic, gastrointestinal or
 CC hepatocellular activity acting as an agonist or antagonist of drug
 CC metabolizing enzyme activity. The invention may be used in the diagnosis
 CC and treatment of disorders associated with decreased or increased
 CC expression or activity of drug metabolizing enzymes. Such disorders
 CC include cancer, cell proliferative disorders, autoimmune/inflammatory,
 CC endocrine, eye, gastrointestinal (including liver disorders) and
 CC metabolic disorders. The present sequence is that of a human drug
 CC metabolizing enzyme (DME) of the invention. Note: This sequence did not
 CC form part of the printed specification but was obtained in electronic
 CC format from EPO.
 XX
 SQ Sequence 618 AA;
 Query Match 84.1%; Score 2588.5; DB 5; Length 618;
 Best Local Similarity 94.4%; Pred. No. 9.6e-233;
 Matches 491; Conservative 1; Mismatches 5; Indels 23; Gaps 1;

QY 31 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPPAPPLGSRFTNPOASPMWDL 90
 DB 92 GPSAEGFORNTRIGMVGKQVTVLGSPPVNVFLGVPPAPPLGSRFTNPOASPMWDL 151
 QY 91 REATSYPNLCLQNSEWLLDQHLKHYPRKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 150
 DB 152 REATSYPNLCLQNSEWLLDQHLKHYPRKGVSEDCLYNTIYAPAHADTGSKLPLYVWF 211
 QY 151 PGGAFTGSAISIFDGSALAAEDVYVWVVOYRLGIGFTTWDQAPGMAFKDOYVALS 210
 DB 212 PGGAFTGSAISIFDGSALAAEDVYVWVVOYRLGIGFTTWDQAPGMAFKDOYVALS 271
 QY 211 WVOKNIEFGDPSSVTIFGESAGAISVSLLISPMAGLPHKAIMESGVATIPYLEAH 270
 DB 272 WVOKNIEFGDPSSVTIFGESAGAISVSLLISPMAGLPHKAIMESGVATIPYLEAH 331
 QY 271 YKSESDLOVVAHFCGNMADSEALRLCLRTKPSKELTLTISQTKSFTTVVDAFPNEBL 330
 DB 332 YKSESDLOVVAHFCGNMADSEALRLCLRTKPSKELTLTISQTKSFTTVVDAFPNEBL 391
 QY 331 DLSQKAFKAIPIISIGVNNHECGFLPMKEAPEILSGSNKSLALHLIQLNIHLI 390
 DB 392 DLSQKAFKAIPIISIGVNNHECGFLPM-----HIPQYLHL 428
 QY 391 VANEFYHDKHSLTEIRDSLLDLGDFVFPVPLITARYHRDAGAVYEFERH 450
 DB 429 VANEFYHDKHSLTEIRDSLLDLGDFVFPVPLITARYHRDAGAVYEFERH 488
 QY 451 TKPAFVKADHDEVAFVFGAFLKGDIVMFGATBEELLSRKMMKWATPARGNP 510
 DB 489 TKPAFVKADHDEVAFVFGAFLKGDIVMFGATBEELLSRKMMKWATPARGNP 548

Qy 511 DLSLMPAYNLTEQYLQDLNMSLGGRLKEPRVDFTWTSTIP 550
Db 549 DLSLMPAYNLTEQYLQDLNMSLGGRLKEPRVDVVTGTP 588

RESULT 10

ID ADB64065 standard; protein; 469 AA.

AC ADB64065;

DT 04-DEC-2003 (first entry)

DE Human protein encoded by clone BRAMH20021910.

Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
cell regeneration; membrane protein; signal transduction-related protein;
transcription-related protein; osteoporosis; neurological disease;
cancer; tumour.

OS Homo sapiens.

PN EP1308459-A2.

PD 07-MAY-2003.

PE 28-MAR-2002; 2002EP-00007401.

PR 05-NOV-2001; 2001JP-00379298.

PR 25-JAN-2002; 2002US-00350978.

PA (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

P1 Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
P1 Yamamoto Y, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
P1 Seki N, Yoshikawa T, Otsuka M, Nagahari K, Maeuho Y;

XX WPI: 2003-450961/43.

DR N-PSDB; ADB62095.

PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.

PS Claim 1; Page: 222pp; English.

The invention discloses a polynucleotide comprising a sequence selected
from 1970 fully defined nucleotide sequences which encode novel

polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
or its partial peptide, an antibody binding to the polypeptide or peptide
of the polynucleotide, immunologically assaying the polypeptide or peptide

peptide of the polynucleotide by contacting the polypeptide or peptide
with the antibody of the encoded protein, and observing the binding
between the two, a transformant carrying the polynucleotide in an

expressible manner and an antisense polynucleotide. The oligonucleotide
is useful as a primer for synthesizing the polynucleotide, or as a probe
for detecting the polynucleotide. The polynucleotides and encoded

proteins are useful as pharmaceutical agents and many disease-related
genes may be included in them, for developing a diagnostic marker or
medicines for regulation of their expression and activity, or as targets

of gene therapy. The genes are involved in tissue and/or cell
regeneration. Membrane proteins, signal transduction-related proteins,
transcription-related proteins, disease-related proteins and genes

encoding them can be used as indicators for diseases (e.g. osteoporosis,
neurological diseases, cancer, tumours. The cDNA may be used to regulate
the activity or expression of the encoded protein to treat diseases. The

sequence presented is a protein of the invention. Note: Some of the
sequence data for this patent is not represented in the printed
specification, but is based on sequence information supplied by the

European Patent Office.

Sequence 469 AA;

Query Match 79.7%; Score 2454; DB 7; Length 469;
Best Local Similarity 99.6%; Pred. No. 2.4e-220;
Matches 467; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 113 MLKVHYPRKGVSEDCLYINIVAPAHADTGSKLPLVWPFPGCAFRTGASIFDGSALAAVE 172
Db 1 MLKVHYPRKGVSEDCLYINIVAPAHADTGSKLPLVWPFPGCAFRTGASIFDGSALAAVE 60
Qy 173 DVLVVVVQYRIGTIGFFFTTMOHA PGWNAFEDQVAAISWVKNIIEFGGDSSTVTI PSES 232
Db 61 DVLVVVVQYRIGTIGFFFTTMOHA PGWNAFEDQVAAISWVKNIIEFGGDSSTVTI PSES 120
Qy 233 AGAIVSSSLISPMAPKGLFHAKIMESGVAIIPYLEAHDEKESBDLQVVAHFCGNNAASDSE 292
Db 121 AGAIVSSSLISPMAPKGLFHAKIMESGVAIIPYLEAHDEKESBDLQVVAHFCGNNAASDSE 180
Qy 293 ALLRCLRTKPSKELLTTSQKTSFTRVVDGAFPPNEPDLISQKAFKAIPSIIGVNSHEC 352
Db 181 ALLRCLRTKPSKELLTTSQKTSFTRVVDGAFPPNEPDLISQKAFKAIPSIIGVNSHEC 240
Qy 353 GPLPMPKEAPRTLSGSNSLSALHLIQLNTLHT PROYLHVANEYFHDKSLTEIRPSLDL 412
Db 241 GPLPMPKEAPRTLSGSNSLSALHLIQLNTLHT PROYLHVANEYFHDKSLTEIRPSLDL 300
Qy 413 LGDVFVVVPAITARVHRDAGAPVYFYEFRHRPOCFEDTKPAFYKADHADEVRFVFGCAF 472
Db 301 LGDVFVVVPAITARVHRDAGAPVYFYEFRHRPOCFEDTKPAFYKADHADEVRFVFGCAF 360
Qy 473 LKGDIVMEGATEEBEKLISRQMKYMATFARTGNPNNGDLSLMPAYNLTEQYLQDLNMS 532
Db 361 LKGDIVMEGATEEBEKLISRQMKYMATFARTGNPNNGDLSLMPAYNLTEQYLQDLNMS 420
Qy 533 LGGRLKEPRVDFTWTSTIPLIISASDMLHSLSLTFLSLQPPFPFCAP 581
Db 421 LGGRLKEPRVDFTWTSTIPLIISASDMLHSLSLTFLSLQPPFPFCAP 469

RESULT 11

ID ADF50145 standard; protein; 542 AA.

AC ADF50145;

DT 12-FEB-2004 (first entry)

DE Cat cauxin protein SEQ ID NO:2.

XX cat; cauxin; cat kidney disease marker; kidney disease.

OS Felis catus.

PN JP2003250575-A.

PD 09-SEP-2003.

PE 04-MAR-2002; 2002JP-00057908.

PR 04-MAR-2002; 2002JP-00057908.

PA (TOHO-) TOHOKU TECHNOARCH KK.

DR WPI: 2004-002277/01.

DR N-PSDB; ADF50144.

PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
PT and for diagnosing cat kidney disease.

PS Claim 3; SEQ ID NO 2; 33pp; Japanese.

The present sequence represents a cat cauxin protein (I) or its salt,
which is cat kidney disease marker. Also described: (1) a partial peptide
(II) of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector

CC (IV) comprising (III); (4) a transformed host (V) comprising (III) or
CC (IV); (5) producing (I) or (II) by culturing (V); (6) an antibody (VI)
CC which comprises specifically with (I) or (II); (7) diagnosing cat kidney
CC disease which involves measuring (I) quantitatively, and where reduction
CC of amount of (I) indicates presence of the disease; (8) a cat kidney
CC disease diagnostic agent comprising (I) labelling agent, a reagent which
CC measures the biological activity of urinary (I) or (VI); and (9) a caution
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (II) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.

SQ Sequence 542 AA;

Query Match	64.5%	Score 1985;	DB 8;	Length 542;
Best Local Similarity	71.7%	Pred. NO. 2.4e-176;		
Matches 373;	Conservative 60;	Mismatches 85;	Indels 2;	Gaps 2

Qy	3	GPAGEGORNTLGTGIMOGQVNTLGS	PVYVNNVFLGVPFAAPRLGSLRPTINPOPASPMNDL	90
Db	22	GPAAADAPVSTRIGLGMWRGKQITVLGS	TVVNNVFLGVPYAAPRLGSLRKPQPRPLPGDF	81
Qy	91	REATSYPNICLQNSWMLLDQHLK	HYHEKFGVSEDCIYLTAYABAHADTGS	KLPVLVWF 15
Db	82	RNATSPYKLCFQDLFWLVS	YOHVLKVRYPKLEASEDCIYLTAYAAHADNSNLP	YVWF 14
Qy	151	PGAFPTGSASIPDGSALAAEDV	LVVVVQYRLGIFGPFPTWDOHAPGNAFKQVVALS	21d
Db	142	PGCAFPMGSASSFDGSALAAEVL	LVTVQYRLGIFGFEDGDEBARNNMALLOQVALT	20
Qy	211	WYQKNIEFFGQDPSSVTIIFGSAGA	ISVSSLTSEPMAGLFFHKAIMESGVALIPYL	EAH 26
Db	202	WYRDNIEFFGQDPBSVTIIFGSAGA	ISVSSLTSPANGLFFHKAIMESGVALITLWRPP	26
Qy	270	DYKESDLOVNAHFCGNNA	SDSEALLRCLRTKPSKELLTISO	KTSFRVYDGAFFPNEP 32
Db	262	GBERKCDLOVLARICGCHAS	DSBALLOCLRAPRPSSEIMDISKUTFS	IPVIDDFFPEDP 32
Qy	330	LDLGSOKAFKALPSIIIGVNNHCE	GCGLTLPMEKAPETLSSGNSSLALHL	IONLHIFPQYTLH 38
Db	322	VALLTOKAFNSVPSIIIGVNNHCE	CAPLL-STEBSSEILGSSNNSLALYL	HTPLNIFQYTLH 38
Qy	390	IVANETVFHDKSLTIRPS	LLDLGLDVFVVYPALITAIRHDAGAPV	FTIEFRHRPOCFE 44
Db	381	LVADHRIPTNKHS	PVEIRIOSFLDLLGLDVLFEVGV	TAHYHDAGAPVYEFYFQHPQCLN 44
Qy	450	DTKPAFVYADHADEVRFVFGA	FLKGDIVMEGATTEBEKLSRQMKMYAT	PARTGPNP 50
Db	441	DTKPAFVYADHSDSEIRFVFGA	FLKGDIVMEGATTEBEKLSRQMKMYAN	FARTGPNP 50
Qy	510	NDLSTWPAYNLTQYLOLDL	MSJGRLKEPRVDFTWTSTI	54
Db	501	BGVPLMPARYTQSEOYTLKDL	LSVSGOYKTLKEOBEVFFMNATTI	54

RESULT 12

AD50147 standard; protein, 542 AA.

AC ADF50147

DT 12-FEB-2004 (first entry)

Cat cauxin protein SEQ ID NO:4.

cat; cat kidney disease marker; kidney disease.

Felis catus.

JP2003250575-A.

PD 09-SEP-2003.
XX
XX PF 04-MAR-2002; 2002JP-00057908.
XX
XX PR 04-MAR-2002; 2002JP-00057908.
XX
XX PA (TOHO-) TOHOKU TECHNORACH KK.
XX
DR WPI; 2004-002277/01.
DR N-PSDB; ADF50146.
XX
XX Novel cauxin protein or its salt, useful as a cat kidney disease marker,
PT and for diagnosing cat kidney disease.
XX
XX Disclosure; SEQ ID NO 4; 33pp; Japanese.
PS

CC This present sequence represents a cat cauxin protein (I) or its salt.
CC Which is cat kidney disease marker. Also described: (1) a partial peptide
CC of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector
CC (IV) comprising (III); (4) a transformed host (V) comprising (III) or
CC (IV); (5) producing (I) or (II) by culturing (V); (6) an antibody (VI)
CC which couples specifically with (I) or (II); (7) diagnosing cat kidney
CC disease which involves measuring (I) quantitatively, and where reduction
CC of amount of (I) indicates presence of the disease; (8) a cat kidney
CC disease diagnostic agent comprising (I) labelling agent; a reagent which
CC measures the biological activity of urinary (I) or (VII); and (9) a cauxin
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as x-ray imaging, ultrasonic imaging,
CC blood testing.

SQ Sequence 542 AA;

Query Match	64.5%	Score 1985;	DB 8;	Length 542;
Best Local Similarity	71.7%;	Pred. No. 2,4e-176;		
Matches 373;	Conservative 60;	Mismatches 85;	Indels 2;	Gaps 2

QY	3	GSASGPOBNTRLMIGOKOVTYVGSVPVNNVLGGDPAAPRLGSLRFTNNPOBAMNDL	90
QY	22	GPADAPRSTRRLGWKQVTVGSVTPVNNVLGGDPAAPRLGSLRFTNNPOBAMNDL	81
QY	91	REATSYPNVLCLQNSEWMLLDQMLKVHPKGVSEBCLYLNITYAPAHADQSKLPELVWF	150
QY	82	RNATSYPLCLQODEMVLVSVOHVLKVRPYPLKLBSEBCLYLNITYAPAHADQSKLPELVWF	141
QY	151	PGAFKTSASAI PGDSALAAVEDLVVVOYRAGIFGPFMTMOAH PGWMAKDDVAALS	210
QY	142	PGAFKTSASAI PGDSALAAVEDLVVVOYRAGIFGPFMTMOAH PGWMAKDDVAALS	201
QY	211	WVOKRIEFGGBESSVTT PGESAGALSVSLLISPAKGLPHKAIMESGVALIPYL-EAH	269
QY	202	WVRDIEEFGGBDPSVTT PGESAGALSVSLLISPAKGLPHKAIMESGVALIPYLMBP	261
QY	270	DYKESBDLOVVAHRCGNNADEBALRCLCTKPSKELLISOKTSPFTVWVGAFPPBP	329
QY	262	GDERKQDQVLAICGCHASDAAALQCRAKRSBELMDISKULPSPIDVDDFFPBP	321
QY	330	LDLSOKAFAPLBIITGVNNHCEGFLPMKEAPELISGNSKSLAHLIOMLITRPVYLH	389
QY	322	VALLTQKFNSSVPSIIGVNNHCEGFL-STERSEIIGGNSRSLAYLHTEFLNTPTVYLH	380
QY	390	LVANEYFHDKSLTEIRDSLDDLLGDVFFVVPVALLTARHNRAGAPVYFEYFRHRPOCFE	449
QY	381	LVADHYFYNNKHSPEVIRDSFLDLGVLFPVPGVLTARHNRAGAPVYFEYFRHRPOCFE	440
QY	450	DTKPAFKADHABEVPVFEAGAFKLDIYMFGAGABEECLSRKMKKMAIPARTGPNNG	509
QY	441	DTKPAFKADHABEVPVFEAGAFKLDIYMFGAGABEECLSRKMKKMAIPARTGPNNG	500
QY	510	NDLSLWPAYNLTVOYLQDLNMSLQGLRLEKPEVPDFTSTI	549

DB 501 EGVPLMPAYTQSEQYLKLDLSVSGKLEQGEVEFFMNTI 540

RESULT 13

ABP61006
ID ABP61006 standard; protein; 356 AA.

XX
AC ABP61006;
XX
DT 10-SEP-2002 (first entry)
XX
DE Novel human protein. SEQ ID 93.

Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian; nootropic; neuroprotective; immunosuppressive; haemostatic; antiinflammatory; cardiant; antidiuretic; vasculide; antithyroid; cerebroprotective; anorectic; metabolic; vaccine; cancer; infection; wound healing disorders; atherosclerosis; Parkinson's disease; Alzheimer's disease; autoimmune disorder; haematopoietic disorder; inflammation; neoplastic disease; nervous system disorder; cardiovascular disorders; pancreatitis; respiratory disorder; hyperproliferation; systemic autoimmune disease; hyper-immunity; developmental abnormality; gastrointestinal ulceration; neuropathy; haematological disease; metabolic disease; sperm dysfunction; thyroid disorder; hypothyroidism; brain damage; colitis; cone photo- transduction deficiency; neurological disease; stroke; angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart; trachea; thymus; lymph node; muscular system; obesity; anorexia; growth abnormality; precocious puberty.

XX Homo sapiens.
OS
XX WO200250105-A1.
PN
XX 27-JUN-2002.
PD
PF 17-DEC-2001; 2001WO-US049232.
PR
XX 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KP, Lal Y,
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX WPI; 2002-508784/54.
DR N-PSDB; AEO86171.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX Claim 1(a); Page 315; 335pp; English.

CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder.
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABP60965-ABP61019 represent novel human
CC proteins of the invention

XX
XX Sequence 356 AA;
SQ

Query Match 57.8%; Score 1780; DB 5; Length 356;
Best Local Similarity 100.0%; Pred. No. 2e-157;
Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ILSPAKGLFKHAIMESGVATIPYLAHDYKSEDLQVVAHPCGNASDSEALRCLRTK 301
DB 17 ILSPAKGLFKHAIMESGVATIPYLAHDYKSEDLQVVAHPCGNASDSEALRCLRTK 76

QY 302 PSKELLTISQKTKSFTRVVDGAFPPNEPLDLSQKAFKAIPIIIVNNHEGFLPMKEA 361
DB 77 PSKELLTISQKTKSFTRVVDGAFPPNEPLDLSQKAFKAIPIIIVNNHEGFLPMKEA 136

QY 362 PEILSGNSKSLALHILQNIHLIPQYLHLVANEYFHDHSLTEIRDSLILDLGDFVFPV 421
DB 137 PEILSGNSKSLALHILQNIHLIPQYLHLVANEYFHDHSLTEIRDSLILDLGDFVFPV 196

QY 422 ALITARYRDAGAPVPEFRRPPOCFEDTKPAFKADHAEVFPVFGAFLKGIWVE 481
DB 197 ALITARYRDAGAPVPEFRRPPOCFEDTKPAFKADHAEVFPVFGAFLKGIWVE 256

QY 482 GATEBEKLLSRKMKKYMATPARTGPNNGNDLSLPAVNLTEQYQLDLMNSLGQRLKEPR 541
DB 257 GATEBEKLLSRKMKKYMATPARTGPNNGNDLSLPAVNLTEQYQLDLMNSLGQRLKEPR 316

QY 542 VDFWTSITPLILSASDMHSPSSITFLSLIQPFPPFCAP 581
DB 317 VDFWTSITPLILSASDMHSPSSITFLSLIQPFPPFCAP 356

RESULT 14

AAE20911
ID AAE20911 standard; protein; 561 AA.

XX
XX AAE20911;
AC
XX 01-JUL-2002 (first entry)
DT
XX
DE Rat carboxylesterase-like enzyme protein.

XX Rat; carboxylesterase-like enzyme; organophosphorus intoxication;
KM osteopathic; gene therapy; osteoporosis; antisense therapy; cytosolic;
KM detoxifying agent; Paget's disease; bone implant degradation; cancer;
KM dental implant; enzyme; gene expression.

XX
XX Rattus sp.
OS
XX WO200206454-A2.
PN
XX 24-JAN-2002.
PD
XX 10-JUL-2001; 2001WO-EP007919.
PF
XX 17-JUL-2000; 2000US-0218564P.
PR
XX (FARB) BAYER AG.
PA
XX
XX Xiao Y;
PI
XX

DR WPI; 2002-195806/25.

XX Novel human carboxylesterase-like enzyme polypeptide, regulators of which
PT are useful for preventing and treating organophosphorus intoxication,
PT cancer and osteoporosis.

PS Disclosure; Fig 8; 92pp; English.

XX The invention relates to a purified human carboxylesterase-like enzyme
CC polypeptide. Carboxylesterase-like enzyme and its DNA are useful for
CC screening for agents which decrease or modulate the activity of
CC carboxylesterase-like enzyme polypeptide. Carboxylesterase-like enzyme is
CC useful for treating a carboxylesterase-like enzyme dysfunction related
CC disease, such as organophosphorus intoxication, cancer and osteoporosis.
CC Compounds that increase the ability of human carboxylesterase-like enzyme
CC to bind to organophosphorus compounds are useful as detoxifying agents.
CC Carboxylesterase-like enzyme agonists and antagonists are useful for
CC treating osteoporosis, Paget's disease and degradation of bone implants,
CC particularly dental implants. Carboxylesterase-like enzyme is useful in
CC diagnostic assays for detecting diseases and abnormalities or
CC susceptibility to diseases or abnormalities related to the presence of
CC mutations in the nucleic acid sequences which encode the enzyme. The
CC coding sequence of carboxylesterase-like enzyme polynucleotide is useful
CC in gene therapy and for generating antisense oligonucleotides or
CC ribozymes which specifically bind to mRNA transcribed from
CC carboxylesterase-like enzyme DNA. These antisense oligonucleotides are
CC useful for modulating carboxylesterase-like enzyme gene expression. The
CC present sequence is rat carboxylesterase-like enzyme protein

XX Sequence 561 AA;

Query Match 41.0%; Score 1262.5; DB 5; Length 561;

Best Local Similarity 48.8%; Pred. No. 1.1e-108;

Matches 275; Conservative 77; Mismatches 179; Indels 33; Gaps 13;

1 MPOG-LTSSAGQWCFLLIQPLLGHQMGKTPSABGQRTKRLGKQVTVLGSVP 59

1 MPRNQLHSLWLVGLL--LLIHVQ-GQDSPES-SPTRTHQVQKLDHVDYDTWAG 56

60 VNVFLGVPPAAPPLGSLRFTNPQAPSPMDNLEATSYNPLCLONSEW----LLDQHLK 115

57 VHTFLGIFPAKPPVPLRPAPEPEPMSGRDASQAMCLQNDILDEGLDMKMT- 115

116 VHYRFGVSECLYNTIAPAHADTGSGLPYLVFPGGAFPTGASIFDGSAALAYEDVL 175

116 --LSISSECLYNTIAPAHADTGSGLPYLVFPGGAFPTGASIFDGSAALAYEDVL 173

176 VVVVYRLGIFGFTTDOHAHGNWAFKDOVAALSWQKNTLFRGDPSSSTTIGESNGA 235

174 VVTIQLRGLVGFSTGDEHARGWNGYIDQVALLRWVQNTLHFGSNFNRVTTFGESAGG 233

236 ISVSSLISPAKGLFHKALMESGVALIPYLEAHDYKSEDLQ--VVAHFCGNMNSDEAL 294

234 TSVSSHVSPMSOGIFGALMESGVALIPDLIS--FTSERVSTTVAKLSCGCEAMDEAL 290

295 LRCRTKPSKRLTLTQKTSFTVVUNGAFFNPBPLDLSKAKALPSTIGVNNHECGF 354

291 VRCIRASGAEILVINKVFKMTPAVVDGEFLPRHPSKLLASEDPHPSPSTIGVTDREYCC 350

355 LLEPM-----KRAEILSGSNKSLALHLIQLIHTIPQYILHVNVEYHDKSLTEIR 406

351 TIPIVMGTALIKR-----LSRENLOAVLKQTAQAQMLPPEGCDLMEYVGNDDSGTLQ 406

407 DSLDLIDGVFPVVALITARYHRDAGAPYTFEFRRPCOFEDTKPAFYKADHADEVRP 466

407 IQYTEMMDGLFVLPALQVAHFOR-SHAPVYFEFQAHAPYFKVAPPHYKADHADEVRP 465

467 VFGAFIKGDIWVEGATEEKLSTRKMKYATFATGPNNGNDLSLWAVNLDTQYLO 526

466 VF-GSFPSSGKLPF---TEERLLSRMMKTMANFARQGNPSEGLPYWALDHDQYLO 521

527 LDIAMSLGQRLEKPRVDFTSTIP 550

DB 522 LDTHPAVDRAKARLQFWTKLP 545

RESULT 15

AAB58981

ID AAB58981 standard; protein; 549 AA.

AC AAB58981;

DT 27-MAR-2001 (first entry)

DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 689.

KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;

KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;

KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;

KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;

KW Addison's disease; allergy; autoimmune haemolytic anaemia;

KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;

KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;

KW cardiovascular disorder; wound healing; neurological disease.

OS Homo sapiens.

PN MO200055173-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US005881.

PR 12-MAR-1999; 99US-0124270P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

XX N-PSDB; AAF21864.

PT New human breast and ovarian cancer associated gene sequences and the

PT polypeptides encoded by these genes, useful in the prevention, treatment

PT and diagnosis of cancer, immune disorders, cardiovascular disorders and

PT neurological diseases.

PS Claim 11; Page 1142-1144; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human

XX proteins AAB58711 - AAB59128. The DNA and protein sequences are

XX associated with breast and ovarian cancer. Included in the invention are

XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the

XX isolation and characterisation of the DNA and protein sequences of the

XX invention. The breast and ovarian cancer associated DNA, protein, agonist

XX or antagonist sequences exhibit cytostatic; immunosuppressive; neotropic;

XX neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;

XX antiinflammatory; antitumor; vulnary; anticonvulsant; antibacterial;

XX antifungal; antiparasitic and cardiac activity. The polynucleotide and

XX protein sequences are used in the diagnosis of cancer, particularly

XX breast and ovarian cancer. The nucleic acid sequences, proteins, agonists

XX and agonists may also be used in the diagnosis, prevention and treatment

XX of immune disorders e.g. Addison's disease, allergies, autoimmune

XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's

XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;

XX cardiovascular disorders such as myocardial ischaemia; wound healing;

XX neurological diseases such as cerebral anoxia and epilepsy; and

XX infectious diseases

XX Sequence 549 AA;

Query Match 39.9%; Score 1230; DB 3; Length 549;

Best Local Similarity 46.7%; Pred. No. 1.2e-105;

Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;

31 GPSAEGPQRNTRLQIGKQVTVLGSPPVNVFLGVPPAAPPLGSLRFTNPQAPSPMDNL 90

Db 16 GDSASPIRTHHTGVLGSLVHVKGANAGVOFLGI PFAPKPLGLRPAPEPPESWGSV 75
Qy 91 REATSYPMLCTON-----SEWLLDQHLKHYHFKGVSEDCLYINTIYAPAHADTGSCLP 145
Db 76 RDGTTHPAMCLODLTAVSEBFL---SQFNMTFPDSMSBEDCLYLSITTPAHSHGSLNP 131
Qy 146 VLVMPGGAFTGASIFDGSALAYEDVLVVVQYRLGIFGPFPTWDOHAPGNWAFKDO 205
Db 132 VVWVHGGALVFGMASTYDGSMLALENVVVIIQYRLGLGFFSTGDKHATGNMGYLDQ 191
Qy 206 VVALSWQKNIEFGGDPSSVTIFGESAGAISVSSLISPMAGLFFHKAIMESGVAIIPY 265
Db 192 VVALRWVQONTIAHFGNPNDRVTIFGESAGTSVSSLVSPISQGLFHGAIMESGVALLPG 251
Qy 266 LEADHYEKSEDL--QVVAHFGCANNASDEBALRCURTSPSKELTLSQTKSFTRVVDGA 323
Db 252 LIA---SSADVISTVVANLSACDQVDSBALVGCIRGKSKERIILAINKPFKMI PGVVDGV 307
Qy 324 FFPNBPDLLSQKAFKAIPSIIGVNNHCGFLLP---MKEAPEILSGSNKSLALHLIQN 379
Db 308 FLFPHPOELLSADFPQVPSTIVGVNNHFGWLIIPKWMRIYDTQKEMDRASQALQKMLT 367
Qy 380 ILHIPPQYILHVANEYFHDKHSLTEIRDSLDDLGDVFPVVVPALITARYHRDAGAPVYFY 439
Db 368 LLMLEPFTGDLREBEYIGDNGDPQTLQAQFQEMWADSMFVILALQVAHF-QCSRAPVYFY 426
Qy 440 EFRHRPOCFEDTKPAFVADHADEVRVFGGAFKGDIVMEGATEEBKLSRRKMKTYA 499
Db 427 EFOHQPSWLKNIRPPHMKADHDDELPPFVF-RSFFCGNYIKF---TEEBEQLSRKMKTYA 482
Qy 500 TEARTGNPNNGDLSIWPAVYNLTEOYLQLDLWMSLGQRLKEPRVDFWTSTIP 550
Db 483 NFARNGNPNNGELPWHPLPQDEBOYLQNLQAPAVGRALKAHRLQPFWKALP 533

Search completed: June 15, 2005, 09:37:45
Job time : 92 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 09:37:21 / Search time 28 Seconds
(without alignments)
1548.968 Million cell updates/sec

Title: US-10-023-515-2
Perfect score: 3079
Sequence: 1 MPQGLTSSASQWCFLLIQP.....PLSSUTFLSLQPPFFCAP 581

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTOS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3079	100.0	581	4	US-10-023-515-2
2	1230	39.9	559	4	US-09-595-682B-28
3	1230	39.9	559	4	US-09-949-016-6426
4	1230	39.9	577	4	US-09-949-016-9670
5	1150.5	37.4	565	4	US-09-595-682B-21
6	1130.5	37.0	566	4	US-09-264-737-2
7	1130.5	36.7	543	4	US-09-595-682B-26
8	1091.5	35.4	584	2	US-08-845-295A-2
9	1091.5	35.4	584	3	US-09-140-933-2
10	1091.5	35.4	584	3	US-09-146-661-2
11	1091.5	35.4	584	3	US-09-150-515-2
12	1034.5	33.6	539	3	US-09-264-737-1
13	942.5	30.6	454	3	US-08-446-100-26
14	942.5	30.6	454	3	US-08-446-100-28
15	942.5	30.6	454	3	US-08-446-100-30
16	942.5	30.6	454	3	US-08-446-100-31
17	940.5	30.5	454	3	US-08-446-100-27
18	940.5	30.5	454	3	US-08-446-100-29
19	871	29.3	574	4	US-10-023-515-4
20	793	25.8	836	4	US-09-949-016-8386
21	786	25.5	823	4	US-09-949-016-6888
22	782	25.4	953	4	US-09-949-016-8387
23	777.5	25.3	843	4	US-09-491-356C-20
24	777	25.2	575	1	US-08-348-920-1
25	775	25.1	823	4	US-09-491-356C-23
26	774	25.1	575	1	US-08-348-920-2
27	772	25.1	848	4	US-09-491-356C-22

28	769	25.0	836	4	US-09-491-356C-21	Sequence 21, Appl
29	765.5	24.9	614	1	US-08-446-100-25	Sequence 25, Appl
30	764.5	24.8	614	3	US-07-732-962A-2	Sequence 2, Appl1
31	764.5	24.8	614	2	US-08-370-156-2	Sequence 2, Appl1
32	764.5	24.8	614	3	US-08-446-100-19	Sequence 19, Appl
33	764.5	24.8	614	3	US-08-446-100-21	Sequence 21, Appl
34	764.5	24.8	614	3	US-08-814-095-2	Sequence 2, Appl1
35	764.5	24.8	614	5	PCT-US92-06106-2	Sequence 2, Appl1
36	764.5	24.8	645	4	US-09-949-016-7063	Sequence 7063, Ap
37	764.5	24.8	645	4	US-09-949-016-7064	Sequence 7064, Ap
38	762.5	24.8	614	3	US-08-446-100-20	Sequence 20, Appl
39	761.5	24.7	614	3	US-08-446-100-23	Sequence 23, Appl
40	760.5	24.7	614	3	US-08-446-100-22	Sequence 22, Appl
41	759.5	24.7	600	2	US-08-370-156-4	Sequence 4, Appl1
42	759.5	24.7	600	3	US-08-814-095-4	Sequence 4, Appl1
43	759.5	24.7	600	3	US-08-975-084-1	Sequence 1, Appl1
44	755.5	24.5	617	2	US-08-370-156-6	Sequence 6, Appl1
45	755.5	24.5	617	3	US-08-814-095-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1									
US-10-023-515-2									
Sequence 2, Application US/10023515									
Patent No. 6664091									
GENERAL INFORMATION:									
APPLICANT: Curtis, Rory A. J.									
TITLE OF INVENTION: Silos-Santiago, Immaculada									
TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE									
FILE REFERENCE: 10448-122001									
CURRENT APPLICATION NUMBER: US/10/023,515									
CURRENT FILING DATE: 2001-12-18									
PRIOR APPLICATION NUMBER: 60/256,369									
PRIOR FILING DATE: 2000-12-18									
PRIOR APPLICATION NUMBER: 60/279,508									
PRIOR FILING DATE: 2001-03-28									
NUMBER OF SEQ ID NOS: 6									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 2									
LENGTH: 581									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-10-023-515-2									
Query Match									
Best Local Similarity 100.0%; Score 3079; DB 4; Length 581;									
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MPQGLTSSASQWCFLLIQPLIGHRQKGTGPSABGPQRNRLGWIQKQVTVLGSPPV	60						
DB	1	MPQGLTSSASQWCFLLIQPLIGHRQKGTGPSABGPQRNRLGWIQKQVTVLGSPPV	60						
QY	61	NVFLGVPFAAPLIGSLRFTNPQASPMNLRBATSYFVLCLQNSFWLLDDHMLKVHPK	120						
DB	61	NVFLGVPFAAPLIGSLRFTNPQASPMNLRBATSYFVLCLQNSFWLLDDHMLKVHPK	120						
QY	121	FGVSEDCILYNIYAPAHADTGSKLPLVLPFGCAFKTGSASIFQGSALAAVEDVLVVVVQ	180						
DB	121	FGVSEDCILYNIYAPAHADTGSKLPLVLPFGCAFKTGSASIFQGSALAAVEDVLVVVVQ	180						
QY	161	YRLGIFGFTTWDDHAPGNWMAFKQVVALSVQKXIEFGDPPSVITFGSAGASIVSS	240						
DB	161	YRLGIFGFTTWDDHAPGNWMAFKQVVALSVQKXIEFGDPPSVITFGSAGASIVSS	240						
QY	241	LILSPMAGLFRHKALMESGVAILPYLEADYKESDLOVVAHFGNNASDBALRLCLRT	300						
DB	241	LILSPMAGLFRHKALMESGVAILPYLEADYKESDLOVVAHFGNNASDBALRLCLRT	300						
QY	301	KPSKELTLISQKTSFTRVVVGAFPPNEPLDLSOKAFKAIPIIIGVNNHCGFLLPNKE	360						
DB	301	KPSKELTLISQKTSFTRVVVGAFPPNEPLDLSOKAFKAIPIIIGVNNHCGFLLPNKE	360						

Db 301 KPSKELLTISQKTSFTRVVDGAFPPNEPDLISQKAFKAPISIIIVNNECGFLPMKE 360
Qy 361 APELISGNSLSALHLIIONLIIHPQYLHLVANEYFHDKSLTEIRDSLLDLGDFEVV 420
Db 361 APELISGNSLSALHLIIONLIIHPQYLHLVANEYFHDKSLTEIRDSLLDLGDFEVV 420
Qy 421 PALITARYHDAGAPVYFYEFRHRPQCFEDTKPAFVADADAEVRFVGGAFKGIWVF 480
Db 421 PALITARYHDAGAPVYFYEFRHRPQCFEDTKPAFVADADAEVRFVGGAFKGIWVF 480
Qy 481 EGATEEKSLSRKKMKWATPARTGNNGNDLSIMPAYNLTQOYLQDLNMSIGORLKEP 540
Db 481 EGATEEKSLSRKKMKWATPARTGNNGNDLSIMPAYNLTQOYLQDLNMSIGORLKEP 540
Qy 541 RVDFTSTPLISASDMLSPISLTFSLLOPFFPCAP 581
Db 541 RVDFTSTPLISASDMLSPISLTFSLLOPFFPCAP 581

RESULT 2
US-09-595-682B-28
Sequence 28, Application US/09595682B
Patent No. 6800483
GENERAL INFORMATION:
APPLICANT: Danke, Mary K.
APPLICANT: Potter, Philip M.
TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
FILE REFERENCE: SJ-0005
CURRENT APPLICATION NUMBER: US/09/595,682B
PRIOR FILING DATE: 2000-01-16
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: PCT/US99/03171
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 559
TYPE: PRT
ORGANISM: Homo sapiens
US-09-595-682B-28

Query Match 39.9%; Score 1230; DB 4; Length 559;
Best Local Similarity 46.7%; Pred. No. 4.8e-125;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;
Qy 31 GPSAEGPQNRTRLGWIOGQVTVLGSPPVNVNPLGVPAAPLGLSIRFTNPQSPMDNL 90
Db 26 GQDSASPIRTHHTGQVLSLVHVGANAAGVOTFLGIPRAKPLGLRPAPEPESWGV 85
Qy 91 REATSYNCLION-----SEWLLDQHLKVHYPKFVSEDCLYLNIYAPAHADTGSKL 145
Db 86 RDGTHPAMCLODLAVASEFL-----SQFNMTPSDMSSEDCLYLSTYTPAHSHSGNLP 141
Qy 146 VLWVPGAFKTSASIFDGSALAAVEDVLYVVOYRLGTFGFTTMDQAPGMARFQD 205
Db 142 VMMVHIGGALVFGMAISLVDSMLALSNVYVYIYRLGVAGFSTGDKAATGMGLDQ 201
Qy 206 VAALSVQKNIIEFGDPSVTIFGESAGAISVSSLISPMACKLPHKAIMESGVAILPY 265
Db 202 VAALRWQONIAHFGNDPRVTIFGESAGTSVSVLSVPSISQGLFHGAIMESGVAILPY 261
Qy 266 LEAHDEKSEDL--QVVAHPCGNNADEALLCRLTPSKELLTISQKTSFTRVVDGA 323
Db 262 LIA-----SSADVISTVVAANLACDQVDEALVGLCRGSKKEILAIKPPKMPGVVDG 317
Qy 324 FFPNEPDLISQKAFKAPISIIIVNNECGFLP---MKEAPEIISGNSKSLALHLION 379
Db 318 FLPRHPQELIASADFOVPISIVGNNEPFGWLLPKWRIYDTQKEMDRASQALQKMLT 377
Qy 380 ILHIPPQYLHLVANEYFHDKSLTEIRDSLLDLGDFEVVPAALITARYHDAGAPVYF 439

Db 378 LMLPPTFGDLIREYIGDNGDPQTLQAQFQEMWADSMFVYPAIQVAHF-QCSAPVYFY 436
Qy 440 EFRHRPQCFEDTKPAFVADADAEVRFVGGAFKGIWVFEGATEEKSLSRKKMKW 499
Db 437 EFQHPQWLKNIIRPPHMKADHGDDELPIVF-RSFGVGNVIKE---TEEBEQLSRKKMKW 492
Qy 500 TPARTGNNGNDLSIMPAYNLTQOYLQDLNMSIGORLKEPVDPMSTIP 550
Db 493 NFARNGNNGGGLPRHWPLFDQEBQYLQNLQPAVGRALKAKHRLQPKKALP 543

RESULT 3
US-09-949-016-6426
Sequence 6426, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6426
LENGTH: 559
TYPE: PRT
ORGANISM: Human
US-09-949-016-6426

Query Match 39.9%; Score 1230; DB 4; Length 559;
Best Local Similarity 46.7%; Pred. No. 4.8e-125;
Matches 248; Conservative 91; Mismatches 168; Indels 24; Gaps 8;
Qy 31 GPSAEGPQNRTRLGWIOGQVTVLGSPPVNVNPLGVPAAPLGLSIRFTNPQSPMDNL 90
Db 26 GQDSASPIRTHHTGQVLSLVHVGANAAGVOTFLGIPRAKPLGLRPAPEPESWGV 85
Qy 91 REATSYNCLION-----SEWLLDQHLKVHYPKFVSEDCLYLNIYAPAHADTGSKL 145
Db 86 RDGTHPAMCLODLAVASEFL-----SQFNMTPSDMSSEDCLYLSTYTPAHSHSGNLP 141
Qy 146 VLWVPGAFKTSASIFDGSALAAVEDVLYVVOYRLGTFGFTTMDQAPGMARFQD 205
Db 142 VMMVHIGGALVFGMAISLVDSMLALSNVYVYIYRLGVAGFSTGDKAATGMGLDQ 201
Qy 206 VAALSVQKNIIEFGDPSVTIFGESAGAISVSSLISPMACKLPHKAIMESGVAILPY 265
Db 202 VAALRWQONIAHFGNDPRVTIFGESAGTSVSVLSVPSISQGLFHGAIMESGVAILPY 261
Qy 266 LEAHDEKSEDL--QVVAHPCGNNADEALLCRLTPSKELLTISQKTSFTRVVDGA 323
Db 262 LIA-----SSADVISTVVAANLACDQVDEALVGLCRGSKKEILAIKPPKMPGVVDG 317
Qy 324 FFPNEPDLISQKAFKAPISIIIVNNECGFLP---MKEAPEIISGNSKSLALHLION 379
Db 318 FLPRHPQELIASADFOVPISIVGNNEPFGWLLPKWRIYDTQKEMDRASQALQKMLT 377
Qy 380 ILHIPPQYLHLVANEYFHDKSLTEIRDSLLDLGDFEVVPAALITARYHDAGAPVYF 439
Db 378 LMLPPTFGDLIREYIGDNGDPQTLQAQFQEMWADSMFVYPAIQVAHF-QCSAPVYFY 436
Qy 440 EFRHRPQCFEDTKPAFVADADAEVRFVGGAFKGIWVFEGATEEKSLSRKKMKW 499
Db 437 EFQHPQWLKNIIRPPHMKADHGDDELPIVF-RSFGVGNVIKE---TEEBEQLSRKKMKW 492

Qy 500 TFAATGPNNGNDLSMPAYNLTEOYLQDLNMSLGRLEKRPVDFWTSTIP 550
Db 493 NFANNGPNNGSLPHWPLFDEBEYLQNLQPAVGRALKARLQFWKALP 543

RESULT 4

US-09-949-016-9670
; Sequence 9670, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9670
; LENGTH: 577
; TYPE: PRF
; ORGANISM: Human
US-09-949-016-9670

Query Match 39.9%; Score 1230; DB 4; Length 577;
Best Local Similarity 46.7%; Pred. No. 5.1e-125;

Matches 24; Conservative 91; Mismatches 166; Indels 24; Gaps 8;

Qy 31 GPEAGGFORTRGLGKQKQVTVLGSPPVNVFLGVPAPPLGSLRPTNDOPASPMNTL 90
Db 44 GQDASPIRTHTQVGLSVHKVAGANAGVQTFIGIPKAPPLGRLPAPBPESWGV 103
Qy 91 REASTYNNLCQN-----SEWLLDDQMLKHYKPKGVSEDCILYNTIYAPHAATGSLP 145
Db 104 RDGTTTHPAMCLODTLVAESBFL---SQFNTTSPDSMSEDCILYITTPHSHGSLP 159
Qy 146 VLWPPGGAFTGTSASIFDGSALAAVEDLVVVVQYRLGFGFTTDOHAPGMMAFMDQ 205
Db 160 VMMVTHGGLVFGNALSIDSMALLAENVVVITQYRLGVGFSTGDKATGMWYLDQ 219
Qy 206 VAALSNVQKNIIEFGDPSSVTIFGESAGALSVSLLSPMAKGLFKKAIMESGVALIPY 265
Db 220 VVALRWQONIAHNGMNDRTTIFGESAGTSVSSLVSPISQGLFGALMESGVALIPG 279
Qy 266 LEANDYKESDL--QVVAHFCGNASDSEALRLCLRTKPSKELLTLSOKTSPTRVVDGA 323
Db 280 LIA-----SSADVISTVANLSACDQVSEALVGLRGSKEEIIAINKPFKMPIDGVVDGV 335
Qy 324 FFPNEPLDLSQKAFKAIPISTIGVNHGCGFLP---MKAPRILSGNSKSLALHLION 379
Db 336 FLPHPELLASADFOVPVSTVGVNNEFGWLTFRKWTIYDQKEMDRBSOALLQKMLT 395
Qy 380 ILHIPPOLYLVANEYFHDKSLTEIRDSLLDLGDVFEVVPALITARYHRDAGAPVYFY 439
Db 396 LLMPLPFPGDLAREYIGDNGDPQLQAQPEMMADSWFVIALQVANH-QCSAPVYFY 454
Qy 440 EPRHRPOCFEBETKAFYKADADAVRPFVFGAFLKGDIVNEBGAITEBKLSRGMKTYA 499
Db 455 EFOHQPSMLKIRIPHHKADHGDDELPFVYF-RSFGGVYIKF---TEEBEQLSRGMKTYA 510
Qy 500 TFAATGPNNGNDLSMPAYNLTEOYLQDLNMSLGRLEKRPVDFWTSTIP 550
Db 511 NFANNGPNNGSLPHWPLFDEBEYLQNLQPAVGRALKARLQFWKALP 561

RESULT 5
US-09-595-682B-21

; Sequence 21, Application US/09595682B

; Patent No. 6800483

; GENERAL INFORMATION:

; APPLICANT: Danks, Mary K.

; APPLICANT: Houghton, Peter J.

; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of

; FILE REFERENCE: SJ-0005

; CURRENT APPLICATION NUMBER: US/09/595,682B

; PRIOR FILING DATE: 2000-01-16

; PRIOR APPLICATION NUMBER: 60/075,258

; PRIOR FILING DATE: 1998-02-19

; PRIOR APPLICATION NUMBER: PCT/US99/03171

; NUMBER OF SEQ. ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 21

; LENGTH: 565

; TYPE: PRF

; ORGANISM: Oryctolagus cuniculus

US-09-595-682B-21

Query Match 37.4%; Score 1150.5; DB 4; Length 565;
Best Local Similarity 43.9%; Pred. No. 2.5e-116;

Matches 252; Conservative 75; Mismatches 180; Indels 67; Gaps 13;

Qy 12 WCFPLIIOPLGRHQRKTPSABGFORNRLGKQKQVTVLGSPPVNVFLGVPAPAP 71
Db 2 WLCLALASLACTAWQH--PSAP-PVVDVHGVKGVSLBGFAPQVAVFLGVPAPAP 58
Qy 72 PLGSLRTPPOPAPSMNDLREASTYPLCLQNSEWMLLDQML-----KVHYP-KFGV 123
Db 59 PLGSLRTPPOPAPSMNHNKTTSTYPMCSODA---VSGHMLSELTNRKENIPLKF-- 112
Qy 124 SEDCLYNTIYAPAHADTGSRLPYLVWPPGAFKTSASIFDGSALAAVEDLVVVVQYRL 183
Db 113 SEDCLYNTIYAPADLTGRGLPVMVWTHGGGLVWVGASTYDGLALSHENVVVVTIQYRL 172
Qy 184 GIGFPTTQOHPAGNKAFFQOVALASWQKNIIEFGDPSSVTIIFESAGALSVSLLIL 243
Db 173 GIMGFSTGDBSHSGNMGHLDVVALRWVDNTANFGDPSGYTIFGESAGGSVSTILL 232
Qy 244 SPMAKGLFKKAIMESGVALIPYLEANDYKESD--LQVVAHFCGNASDSEALRLCLRTK 301
Db 233 SPLTNLFHRAISBSGVALSSL---FRKNTSLAEKILAECKTTTSAWVHCLRQK 288
Qy 302 PSKELL--TISOK-----TKSFTRVVDGAFPPNEPLDLSQKAFKAIPIPSIIG 346
Db 289 TEBELMEVTILKMKFMALDVGDPKENTAFITTYIDGVLTKAPAEILAEKKYNNLPYMWG 348
Qy 347 VNHGCGFLPM-----KEAPEILSGNSKSLALHLIONILHIPPOLYLVVA 392
Db 349 INQOEFMIIPMOGLGYPLSEGLDKQKTATELWKS-----YPIVNVSKELTPTVAT 399
Qy 393 NEYFHDGSLTEIRDSLLDLGDVFEVVPALITARYHRDAGAPVYFEPHRRPOCFEDTK 452
Db 400 EKYTGTDTPVKKXDLFLMDLADLFGVPSVNVVARRHRDGAFTMYEYIRFSFSDMK 459
Qy 453 PAFYKADADAVRPFVFGAFLKGDIVNEBGAITEBKLSRGMKTYATFARTGPNNGNDL 512
Db 460 PKTYIGDHGDEIPSVLGAPFLK-----EGATEBEIRLISGMWKKYANFANNGPNNGEGL 513
Qy 513 SLMPAYNLTEOYLQDLNMSLGRLEKRPVDFWT 546
Db 514 POWPAYDYKEGYLQIGATTQAAOKLKDKEVAFWT 547

RESULT 6
US-09-264-737-2
; Sequence 2, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:

```

; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ruff, Thomas G.
; TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
; TITLE OF INVENTION: Expression of Esterase Enzymes
; FILE REFERENCE: 38-21(10551) RUES Pyridine Tolerance
; CURRENT APPLICATION NUMBER: US/09/264,737A
; EARLIER FILING DATE: 1999-03-09
; EARLIER APPLICATION NUMBER: 60/077,377
; EARLIER FILING DATE: 1998-03-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Rabbit
; US-09-264-737-2

```

```

Query Match      37.0%; Score 1138.5; DB 3; Length 566;
Best Local Similarity 43.6%; Pred. No. 5.1e-115;
Matches 250; Conservative 75; Mismatches 176; Indels 73; Gaps 14;

```

```

QY 12 WCFLLIQLPLGHRQKTPSAGPQRTRLGWIQKQVTVLGSVPVNVFLGVPPAP 71
DB 9 WLF-----LAACIATMGGH--PSAP-PVVDIVHGKVLGKFSLEGPAQVAVFLGVPPAP 59
QY 72 PLGSLRFTNPQAPSPMDNLRATSYRNLCTONSEWMLLDQHML-----KVHYP-KFGV 123
DB 60 PLGSLRFAPQAPASPSHVKNTTSYRPMCSQDA-----VSGHMLSELFTNRKENIPLKF-- 113
QY 124 SEDCLYNTIYAPAHADTGSKLPLYWPPGGAFTGASIFDGSALAAVEDVLVVVVQYRL 183
DB 114 SEDCLYNTIYTPADLTGKGRLLPVMVWHGGGLWVGASTYDGLALSAHENVVVVTIQYRL 173
QY 184 GIGFFFTTWDQAPGNMAFKDOVALASWQKNIIEFGGDPSSVTITFGESAGASVSLL 243
DB 174 GIMGFSTGDEHSGKNGHLDQVRLRWQDNIAFGDPSGVTITFGESAGGQSVSILL 233
QY 244 SPMAKGLFPAKAMEGVAIIPLYLAHDYKESD--LQVVAHFCGNNASDEALRCLRTK 301
DB 234 SPLTKMLFRAIISSEGVALLSL-----FRKNTKSLAEKIAIEAGCKTTTSAVMVHCLRQK 289
QY 302 PSKELL--TLQSK-----TKSFTRVNDGAFPPNEBLLDSOKAFKAIPSTIG 346
DB 290 TEEBLMEVTLKMKFPAALDVGDPKENTAFLLTVIVGVLIPKAPABITLAEKKYNNMLPYWG 349
QY 347 VNNHCEGFLPM-----KEAPILSGSNKSLALHLIQLIHLIPQYLHLVA 392
DB 350 INQOEFGWIIIPQMLGYPLSEKGLDQKATATLWKS-----YPIVNVSKELTPVAT 400
QY 393 NEYFHDKSLTEIRDSLLDLGDVFFVVPALITARYHRDAGAPVYFEFRHRPQCFEDTK 452
DB 401 EKVIGGTDPPVKKDLFLDMLADLLFGVSVVVAHRHDAGAPTYMEYRVRPSSDMR 460
QY 453 PAFVKADHADEVRFVFGAFLKGDIVMFEAGATEEEKLSRKMKMYATFARTGNNGNDL 512
DB 461 PKTVIGDHDDELFSVLGAPFLK-----EGATEEIRIKLSKVMKYMANFARNGNNGGL 514
QY 513 SLMPAYNLTQYLTQDLNMSLQRLKEPRVFWT 546
DB 515 PQMPAYDYKEGYLQIGATTAQAQKLKDXEVAFWT 548

```

```

RESULT 7
; US-09-595-682B-26
; Sequence 26, Application US/09595602B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SJ-0005

```

```

; CURRENT APPLICATION NUMBER: US/09/595,682B
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 26
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; US-09-595-682B-26

```

```

Query Match      36.7%; Score 1130.5; DB 4; Length 543;
Best Local Similarity 43.7%; Pred. No. 3.5e-114;
Matches 249; Conservative 75; Mismatches 179; Indels 67; Gaps 13;

```

```

QY 12 WCFLLIQLPLGHRQKTPSAGPQRTRLGWIQKQVTVLGSVPVNVFLGVPPAP 71
DB 2 WLCALALASLAACTAMGH--PSAP-PVVDIVHGKVLGKFSLEGPAQVAVFLGVPPAP 58
QY 72 PLGSLRFTNPQAPSPMDNLRATSYRNLCTONSEWMLLDQHML-----KVHYP-KFGV 123
DB 59 PLGSLRFAPQAPASPSHVKNTTSYRPMCSQDA-----VSGHMLSELFTNRKENIPLKF-- 112
QY 124 SEDCLYNTIYAPAHADTGSKLPLYWPPGGAFTGASIFDGSALAAVEDVLVVVVQYRL 183
DB 113 SEDCLYNTIYTPADLTGKGRLLPVMVWHGGGLWVGASTYDGLALSAHENVVVVTIQYRL 172
QY 184 GIGFFFTTWDQAPGNMAFKDOVALASWQKNIIEFGGDPSSVTITFGESAGASVSLL 243
DB 173 GIMGFSTGDEHSGKNGHLDQVRLRWQDNIAFGDPSGVTITFGESAGGQSVSILL 232
QY 244 SPMAKGLFPAKAMEGVAIIPLYLAHDYKESD--LQVVAHFCGNNASDEALRCLRTK 301
DB 233 SPLTKMLFRAIISSEGVALLSL-----FRKNTKSLAEKIAIEAGCKTTTSAVMVHCLRQK 288
QY 302 PSKELL--TLQSK-----TKSFTRVNDGAFPPNEBLLDSOKAFKAIPSTIG 346
DB 289 TEEBLMEVTLKMKFPAALDVGDPKENTAFLLTVIVGVLIPKAPABITLAEKKYNNMLPYWG 348
QY 347 VNNHCEGFLPM-----KEAPILSGSNKSLALHLIQLIHLIPQYLHLVA 392
DB 349 INQOEFGWIIIPQMLGYPLSEKGLDQKATATLWKS-----YPIVNVSKELTPVAT 399
QY 393 NEYFHDKSLTEIRDSLLDLGDVFFVVPALITARYHRDAGAPVYFEFRHRPQCFEDTK 452
DB 400 EKVIGGTDPPVKKDLFLDMLADLLFGVSVVVAHRHDAGAPTYMEYRVRPSSDMR 459
QY 453 PAFVKADHADEVRFVFGAFLKGDIVMFEAGATEEEKLSRKMKMYATFARTGNNGNDL 512
DB 460 PKTVIGDHDDELFSVLGAPFLK-----EGATEEIRIKLSKVMKYMANFARNGNNGGL 513
QY 513 SLMPAYNLTQYLTQDLNMSLQRLKEPRVFWT 542
DB 514 PQMPAYDYKEGYLQIGATTAQAQKLKDXEVAFWT 543

```

```

RESULT 8
; US-08-845-295A-2
; Sequence 2, Application US/08845295A
; Patent No. 5817490
; GENERAL INFORMATION:
; APPLICANT: Hubbs, John C.
; TITLE OF INVENTION: Enzymatic Process for the Manufacture of
; TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Chemical Company
; STREET: P. O. Box 511
; CITY: Kingsport

```


STATE: Tennessee
COUNTRY: USA
ZIP: 37662-5075
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,295A
FILING DATE: 25-April-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,879
FILING DATE: 17-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Cheryl J. Tubach
REGISTRATION NUMBER: 38,346
REFERENCE/DOCKET NUMBER: 70432
TELECOMMUNICATION INFORMATION:
TELEPHONE: 423-229-6189
TELEFAX: 423-229-1239
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-845-295A-2

Query Match 35.4%; Score 1091.5; DB 2; Length 584;
Best Local Similarity 43.8%; Pred. No. 7.4e-110;

Matches 259; Conservative 63; Mismatches 191; Indels 79; Gaps 17;

12 WCFPLIOLPLIGHOMKGTGSAAGPQNRRLGWIQKQVTV--LGSPVPPNVFLGVPPA 69
2 WILPLVLITLSSASATW--AGOPASPPVVDTAQGRVLGKTVSLGGLAFTQPAVFLGVPPA 59
70 APPGISLRFNTPOPASPMNDLREATSYNLCLONS--EWLLD-----OHMLKVHPK 120
60 KPLGSLRFAPPOAPAEPMSPVKNTTSYPMCCQDVPVQMTSDLEFTNFGKERLTLEF-- 117
121 FGVEDCLYINITYAPAHADTGSKLPLVWPFPGAFTKGSASIFDGSALAAYED--VLVVV 178
118 ---SEDCLYINITYPADLTTRGRPLPVWVWIHGGGLVVGAGPMYGVVLAHENEFTVVVVA 174
179 VQYRLGIFGFPTTWDQAHFGWMAFKDQVVALSWVQKNIFFGGDSSVTITGES--AGAI 236
175 IQYRLGIFGFPTTWDQAHFGWMAFKDQVVALSWVQKNIFFGGDSSVTITGES--AGAI 236
237 SVSSILSPMAKGLFHAKAIMESGVAIIPYLEAHDIKESDQLQVAHFCGNNAASEALLR 296
235 SVSVLVLSPLAKNLFHRAISSGVALTVLVKMKMAAKQIAYLAAGCKTTTSAVFTFVH 294
237 CLRTKPSKELLTLISQKTSFT-----RVVDGAFPPNBPDLILSOK--AFK 339
297 CLRTKPSKELLTLISQKTSFT-----RVVDGAFPPNBPDLILSOK--AFK 339
295 CLRQKSEBELDLTLKKKFLTLDFHGDQRBESHPLPLTVVDGVLLPKMBEETILAEKDFTFN 354
340 AIPSLIIGVNNHECGFLRP-----MKBAPEILSGSNKSLAHLIONTLHIPPQ 386
355 TVPIYVIGINKQEPGLLPTVMGPPLSBGKLDQKNTLSLWKS-----YPLANI-----PE 404
387 YLHLVAAEFYDHA-----SLTEIRDSLDDLGDFVFNVPALITKRYHDKAGAPYTFEF 441
405 ELTPVAT--PTDKYLGSTDDPVKKKDLFLDLMGDVVPFVPSVTVAHQHDDGAPTYMEF 462
442 RHRPCCPED--TKPAFYADNADEVRFPVGGAFLLGDTVMFPGATEEKKLSRKKMKYTA 499
463 QYRBSFSSDKTKKPTVYIGDHGDELFSVFGPPLNGD-----ABEEVSLSKTYWKEWA 516
500 TPARTGNNGNDLSLMP--AYNLTQOYLQDLNMSLAGRLKEPRVDFTSTI 549
517 NFASGNNGSGLPHMPFTWYDQEGGYLQIGVNTQAARLKGEBVAFVNDLL 568

RESULT 9

US-09-140-933-2

; Sequence 2, Application US/09140933

; Patent No. 6022719

; GENERAL INFORMATION:

; APPLICANT: Hubbs, John C.

; TITLE OF INVENTION: Enzymatic Process for the Manufacture of

; TITLE OF INVENTION: Acarboxic Acid, 2-Keto-L-Gulonic Acid, and Esters of

; TITLE OF INVENTION: 2-Keto-L-Gulonic Acid

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESS: Eastman Chemical Company

; STREET: P.O. Box 511

; CITY: Kingsport

; STATE: Tennessee

; COUNTRY: USA

; ZIP: 37662-5075

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch disk

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows 95

; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/140,933

; FILING DATE: 27-August-98

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/017,879; 08/845,295

; FILING DATE: 17-May-96; 25-April-97

; ATTORNEY/AGENT INFORMATION:

; NAME: Cheryl J. Tubach

; REGISTRATION NUMBER: 38,346

; REFERENCE/DOCKET NUMBER: 70432

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 423-229-6189

; TELEFAX: 423-229-1239

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 584 amino acids

; TYPE: Amino Acid

; TOPOLOGY: Linear

; MOLECULE TYPE: protein

US-09-140-933-2

Query Match 35.4%; Score 1091.5; DB 3; Length 584;
Best Local Similarity 43.8%; Pred. No. 7.4e-110;

Matches 259; Conservative 63; Mismatches 191; Indels 79; Gaps 17;

12 WCFPLIOLPLIGHOMKGTGSAAGPQNRRLGWIQKQVTV--LGSPVPPNVFLGVPPA 69
2 WILPLVLITLSSASATW--AGOPASPPVVDTAQGRVLGKTVSLGGLAFTQPAVFLGVPPA 59
70 APPGISLRFNTPOPASPMNDLREATSYNLCLONS--EWLLD-----OHMLKVHPK 120
60 KPLGSLRFAPPOAPAEPMSPVKNTTSYPMCCQDVPVQMTSDLEFTNFGKERLTLEF-- 117
121 FGVEDCLYINITYAPAHADTGSKLPLVWPFPGAFTKGSASIFDGSALAAYED--VLVVV 178
118 ---SEDCLYINITYPADLTTRGRPLPVWVWIHGGGLVVGAGPMYGVVLAHENEFTVVVVA 174
179 VQYRLGIFGFPTTWDQAHFGWMAFKDQVVALSWVQKNIFFGGDSSVTITGES--AGAI 236
175 IQYRLGIFGFPTTWDQAHFGWMAFKDQVVALSWVQKNIFFGGDSSVTITGES--AGAI 236
237 SVSSILSPMAKGLFHAKAIMESGVAIIPYLEAHDIKESDQLQVAHFCGNNAASEALLR 296
235 SVSVLVLSPLAKNLFHRAISSGVALTVLVKMKMAAKQIAYLAAGCKTTTSAVFTFVH 294
237 CLRTKPSKELLTLISQKTSFT-----RVVDGAFPPNBPDLILSOK--AFK 339
297 CLRTKPSKELLTLISQKTSFT-----RVVDGAFPPNBPDLILSOK--AFK 339
295 CLRQKSEBELDLTLKKKFLTLDFHGDQRBESHPLPLTVVDGVLLPKMBEETILAEKDFTFN 354

TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-09-150-515-2

Query Match 35.4%; Score 1091.5; DB 3; Length 584;

Best Local Similarity 43.8%; Pred. No. 7,4e-110;
Matches 259; Conservative 63; Mismatches 191; Indels 79; Gaps 17;

```

QY 12 WCFPIILQPLGHRWGKTGPSAEGQRTKLGMIQKQVTV--LGSPVNVVFLGVPRA.69
DB 2 WLPPIVLTSLSSASAT--AGOPASPPVVDVTAQGRVLGKVSLEGLAFTQPVAVFLGVPRA.59
QY 70 APLSLRFTNPQAPSPMDNLEATSPYLCLONS--EWLLD-----GMLKHVYK.120
DB 60 KPPSLRPAQAPQAPSPMPSFVAKTTSYPMCCODPVVEQMTSDLPFTFGKRLTLER--117
QY 121 FGVSDECLYLNTPAPAHADTGSKLPLYVWFGGAFKTSASIFDGSALAAYED--VLVVV.178
DB 118 ---SDCLYLNITYPADLTKRGRLPYMWVVIHGGGLVLGGAPMYDGVVLAHNFVTVVVA.174
QY 179 VQYRIGTIGFTTQOHAPGNWAFKDQVAAALSMVQKNIFFGSDPSSTVTFGES--AGAI.236
DB 175 IQYRLGIGFFSTGDESHKGNMGLDQVAAALTMVQENIANFGSDPSSTVTFGESFTAGGE.234
QY 237 SVSLLISPMAGLPHKAIMESGVAIIPYLAHDYKESDQLQVVAHPCGNNASDSEALLR.296
DB 235 SVSIVLSPPLAKNLPHRAISBSGVALTVAVLRKMAAKQJAVLAGCTTTTSANVFTVH.294
QY 297 CLRTKPSKELLTLQKTKSFT-----RVVDGAFPPNEPLDLSQK--AFK.339
DB 295 CLRQKSEDELDLTLKMKFLTLDPHQDQRESHPLPTVVDGVLPRKMPBEILAEKDFTFN.354
QY 340 AIPSTIGVNNHCGFLP-----MKAPBEILSGNSKSLAHLIQLNIHIIHPQ.386
DB 355 TVPYIVGINKRFGMLPTMGFPLSEGLDQKTSLSLMS-----YPIINI-----PE.404
QY 367 YLHVAANEYFHDKH-----SLTEIRDSLLDLGADVFPVVPALITARYHNDAGAPYFYEF.441
DB 405 ELTPVAT--FTDKYAGTDDPVKKDLFLDLGDDVFGVSVTVARQHRDAGAPTYMTIEF.462
QY 442 RHPOCFED--TKPAFVKAHADAEVAFVFGAFLKGDIVMEFEGATEEERKLSRKMKTYA.499
DB 463 QYRPFSSDKPTKPKTVIGDHDEIFSVRGPFLKGD-----APBEVLSKTKMKFYA.516
QY 500 TPARTGNPNDLSLP--AYNLTEQYIQLDINMSLGORLKEPRVDFTSTI.549
DB 517 NFARSGNPNGBGLPMMPFTMYDQEGYLIQIVNTQAAKRLKGEVAFVNDLL.568

```

RESULT 12

US-09-264-737-1

/ Sequence 1, Application US/09264737A

/ Patent No. 6107549

/ GENERAL INFORMATION:

/ APPLICANT: Beng, Paul C.C.

/ APPLICANT: Ruff, Thomas G.

/ TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via

/ FILE REFERENCE: 38-21(1051) RLE3 Pyridine Tolerance

/ CURRENT FILING DATE: 1999-03-09

/ EARLIER APPLICATION NUMBER: 60/077,377

/ NUMBER OF SEQ ID NOS: 11

/ SOFTWARE: Patent In Ver. 2.0

/ SEQ ID NO 1

/ LENGTH: 539

/ TYPE: PRT

/ ORGANISM: Rabbit

/ US-09-264-737-1

Query Match 33.6%; Score 1034.5; DB 3; Length 539;

Best Local Similarity 42.2%; Pred. No. 1.1e-103;
Matches 234; Conservative 71; Mismatches 176; Indels 73; Gaps 13;

```

QY 32 PSABGPQNTRLGMIQKQVTVLGSPPVNVVFLGVPAPAPLGSIRFNPNQASPMNLR.91
DB 2 PSAP--PVDIVYAGKVLGKVSLEGAQPVAVFLGVFPKPPGSLRFAPQPAESWSHVK.60
QY 92 EATSPNLCQSEWLLDOHML-----KVHY-PKGVSEDCYLNTPAPAHADTGSK.143
DB 61 NNTTSYPMCSSA---VSGHMLSELFTNRKENITLKP--SDCLYLNITYPADLTKGR.114
QY 144 LPVLVWFGGAFKTSASIFDGSALAAYEDVLVVVQYRLGIGFPTTWDQAPGNMAFK.203
DB 115 LPVWVVIHGGGLMVGASTYDGLALSAHENVVVTVIQYRLGIGGFGFNIDE-----L.166
QY 204 DQVAAALSMVQKNIFFGSDPSSTVTFGESAGALSVSLLISPMAGLPHKAIMESGVAI.263
DB 167 FLVAVNRVVDIANFGDPSGVTTFGESAGQSVSILLSPITNLPHRAISBSGVAL.226
QY 264 PYLAHDYKESD--LQVVAHPCGNNASDSEALLRCLRTKPSKELL--TLQK-----312
DB 227 SSL-----FRKTKSLAEKIALTAGCTTTSAMVHCLQKTEBELMEVTLKMKFMAOLDV.282
QY 313 -----TKSPTRVVDGAFPPNEPLDLSQKAFKALPSIIGVNNHCGFLP-----358
DB 283 GDPKENTAFLTVDIGVLLPKAPAEIYEKKYKMLPYWGINQGFWMIIPMQMLGYPLS.342
QY 359 -----KEAPBEILSGNSKSLAHLIQLNIHIIHPYQYLAHVAANEYFHDKSLTEIRPSL.412
DB 343 BGKLDQKATLTLMS-----YPIVNSKELLTPATEKYLGTDDPVKKDLFLDM.393
QY 413 LGDVFVPVVALITARYHNDAGAPYFYEFRRHPOCFEDTKPAFVKAHADAEVAFVFGAF.472
DB 394 LADLFGVSVVVAHHRDAGAPTYMYRTRPSSSDMRPVTVIGDDEIFSLGAPF.453
QY 473 LKGDIVMEFEGATEEERKLSRKMKTYAFARTGNPNDLSLPAYNLTEQYIQLDINMS.532
DB 454 LK-----EGATEEIRKLSRKMKTYAFARTGNPNDLSLPAYNLTEQYIQLDINMS.507
QY 533 LGORLKEPRVDFT.546
DB 508 AAKKLKXKVAFWT.521

```

RESULT 13

US-08-446-100-26

/ Sequence 26, Application US/08446100

/ Patent No. 6001625

/ GENERAL INFORMATION:

/ APPLICANT: Broomfield, Clarence A

/ APPLICANT: Millard, Charles B

/ APPLICANT: Lockridge, Oksana

/ TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases

/ NUMBER OF SEQUENCES: 31

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: Hendricks and Assoc.

/ STREET: 9669 A Main Street, P.O. Box 2509

/ CITY: Fairfax

/ STATE: VA

/ COUNTRY: US

/ ZIP: 22031

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patent In Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/446,100

/ FILING DATE: 19-MAY-1995

/ CLASSIFICATION: 435

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Hendricks, Glenna

/ REGISTRATION NUMBER: 32,535

```
/
/ REFERENCE/DOCKET NUMBER: broomfield
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 425-4250
/ TELEFAX: (703) 425-2767
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 454 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ANTI-SENSE: YES
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: human esterases
/
US-08-446-100-26

Query Match          30.6%; Score 942.5; DB 3; Length 454;
Best Local Similarity 44.9%; Pred. No. 9.5e-94;
Matches 200; Conservative 64; Mismatches 150; Indels 31; Gaps 6;

QY 125 EDCLYNTIYPAHADGSKLPVLVWFPFGAGFKTGSASIFDGSALAAVEDLVVVVQYRLG 184
DB 1 EHCLYNTIYTPADLTQKQRLPVMVWIHGSGLMVGAASYDGLAALAHENVVVVTIYRLG 60
QY 185 IFGFPTTDOHAPGNWAFPDQVAALSWYQKNIFFPGDPSVTITFGESAGATSVSSLLIS 244
DB 61 IWGFSTGDHSHKNGHLDQVAALRWQDNINASTFGNPGSVTITFGESAGSBSVSVLVLS 120
QY 245 PMAKGLFHKAIMESGVAAIIPYLEAHDYER--SEDLQVVAHFCCGNNAUSDSEALLRCLRTKP 302
DB 121 PLAKNLFHRAISSEGVALTSLVKKGDVKKPLAEQIAITA--GCKTTTSAAVHCHLRQKT 177
QY 303 SKELLTISOK-----TKSFTRVVDGAFPNRBPDLISQAKFAIPBIIIGV 347
DB 178 EEBLETTTLKIGNSYLMYRTQRESTLLGTVIDMLLKTPEBLQRENFHTVPYMWGI 237
QY 348 NNHECGFLPMKEAPEILSGS--NKSIALHLION---ILHIPOYLHLVANEYFHDKHSL 402
DB 238 NKQEPGMILPMQMSYPLBSGQLODKTMSILGSIPLPALAKELIPATEKTLGGTDDT 297
QY 403 TEIRDSLIDLIDGVFVVVPALITARYHRDAGAPVYFEFRHRPOCFEDTKPAFYKADHAD 462
DB 298 VKKCDLILDLIADVMFGVPSYIVARNHRDAGAPTYMFEFYRPSFSQMKPKTVIGDHG 357
QY 463 EVRFVFGAFLKGDIMVFGATTEEEKLSRKMKYATARCPNPGNDLSLMPAYNLTE 522
DB 358 ELFSVFGAFLK-----EGASEEBIRLSKMKYKFWANFARNPNKGGLPHMPEYNOKE 411
QY 523 QYQLDINMSLQRLKEPRVDFTWS 547
DB 412 GYLQIGANTQAQKLDKXEVAFWTN 436

RESULT 14
US-08-446-100-28
/ Sequence 28, Application US/08446100
/ Patent No. 6001625
/ GENERAL INFORMATION:
/ APPLICANT: Broomfield, Clarence A
/ APPLICANT: Millard, Charles B
/ APPLICANT: Lockridge, Okeana
/ TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Hendricks and Assoc.
/ STREET: 9669 A Main Street, P.O. Box 2509
/ CITY: Fairfax
/ STATE: VA
/ COUNTRY: US
/ ZIP: 22031
/ COMPUTER READABLE FORM:
```

```
/
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/446,100
/ FILING DATE: 19-MAY-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hendricks, Glenna
/ REGISTRATION NUMBER: 32,535
/ REFERENCE/DOCKET NUMBER: broomfield
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 425-4250
/ TELEFAX: (703) 425-2767
/ INFORMATION FOR SEQ ID NO: 28:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 454 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHEICAL: YES
/ ANTI-SENSE: YES
/ FRAGMENT TYPE: N-terminal
/ ORIGINAL SOURCE:
/ ORGANISM: human esterases
/
US-08-446-100-28

Query Match          30.6%; Score 942.5; DB 3; Length 454;
Best Local Similarity 44.9%; Pred. No. 9.5e-94;
Matches 200; Conservative 65; Mismatches 149; Indels 31; Gaps 6;

QY 125 EDCLYNTIYPAHADGSKLPVLVWFPFGAGFKTGSASIFDGSALAAVEDLVVVVQYRLG 184
DB 1 EHCLYNTIYTPADLTQKQRLPVMVWIHGHLBVGAASTYDGLAALAHENVVVVTIYRLG 60
QY 185 IFGFPTTDOHAPGNWAFKQVAALSWYQKNIFFPGDPSVTITFGESAGATSVSSLLIS 244
DB 61 IWGFSTGDHSHKNGHLDQVAALRWQDNINASTFGNPGSVTITFGESAGSBSVSVLVLS 120
QY 245 PMAKGLFHKAIMESGVAAIIPYLEAHDYER--SEDLQVVAHFCCGNNAUSDSEALLRCLRTKP 302
DB 121 PLAKNLFHRAISSEGVALTSLVKKGDVKKPLAEQIAITA--GCKTTTSAAVHCHLRQKT 177
QY 303 SKELLTISOK-----TKSFTRVVDGAFPNRBPDLISQAKFAIPBIIIGV 347
DB 178 EEBLETTTLKIGNSYLMYRTQRESTLLGTVIDMLLKTPEBLQRENFHTVPYMWGI 237
QY 348 NNHECGFLPMKEAPEILSGS--NKSIALHLION---ILHIPOYLHLVANEYFHDKHSL 402
DB 238 NKQEPGMILPMQMSYPLBSGQLODKTMSILGSIPLPALAKELIPATEKTLGGTDDT 297
QY 403 TEIRDSLIDLIDGVFVVVPALITARYHRDAGAPVYFEFRHRPOCFEDTKPAFYKADHAD 462
DB 298 VKKCDLILDLIADVMFGVPSYIVARNHRDAGAPTYMFEFYRPSFSQMKPKTVIGDHG 357
QY 463 EVRFVFGAFLKGDIMVFGATTEEEKLSRKMKYATARCPNPGNDLSLMPAYNLTE 522
DB 358 ELFSVFGAFLK-----EGASEEBIRLSKMKYKFWANFARNPNKGGLPHMPEYNOKE 411
QY 523 QYQLDINMSLQRLKEPRVDFTWS 547
DB 412 GYLQIGANTQAQKLDKXEVAFWTN 436

RESULT 15
US-08-446-100-30
/ Sequence 30, Application US/08446100
/ Patent No. 6001625
/ GENERAL INFORMATION:
/ APPLICANT: Broomfield, Clarence A
/ APPLICANT: Millard, Charles B
```

APPLICANT: Lockridge, Oksana
 TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hendricks and Assoc.
 STREET: 9669 A Main Street, P.O. Box 2509
 CITY: Fairfax
 STATE: VA
 COUNTRY: US
 ZIP: 22031
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,100
 FILING DATE: 19-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Hendricks, Glenna
 REGISTRATION NUMBER: 32,535
 REFERENCE/DOCKET NUMBER: Bloomfield
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 425-4250
 TELEFAX: (703) 425-2767
 INFORMATION FOR SEQ ID NO: 30:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 454 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: YES
 ANTI-SENSE: YES
 FRAGMENT TYPE: N-terminal
 ORIGINAL SOURCE:
 ORGANISM: human esterases
 US-08-446-100-30
 Query Match 30.6%; Score 942.5; DB 3; Length 454;
 Best Local Similarity 44.9%; Pred. No. 9.5e-94;
 Matches 200; Conservative 64; Mismatches 150; Indels 31; Gaps 6;
 QY 125 EDCIYINITYAPAHADTSGSKLPYLVWFGGAKTGSASIFDGSALAAYEDVLVWVQYRLG 184
 DB 1 EHCYLYNITYPADLTKKRLPVMWVHGGGLMVGAASTYDGLALAHENVVVYTIQYRLG 60
 QY 185 IFGFFTTDOHAPGWMAPKDOVALSWQKNIFFPGDPSVTIFGESAGAI SVSSLTLS 244
 DB 61 IWFSTSGDHSRGNWGHLDVAPALRWQDNINASFGNPGSVTIFGESAGBSVSVLVLS 120
 QY 245 PMAKGLPKAIAMESGVALLPYLEADYEK--SEDLQVVAHFCGNNASDSEALLRCLRTKP 302
 DB 121 PLANKLTFRAISESGVALTSVLKKGDYKPLAEQAIAT--GCCTTSAAWVHCLRQKT 177
 QY 303 SKELLTTSQK-----TKSFTRVVDGAPFPNEPLDLSQAKAIPSIIGV 347
 DB 178 EBELETTTLKIGNSYLTWYRETORHSTLGTVIDGMLTKTPEBLORERNFHTVPYWG 237
 QY 348 NNHECGFLLPKKEAPELISGS--NKSIALHLION---ILHI PPOYLHVANEYFHDKSL 402
 DB 238 NKQFGWLIIFQALSYPSSEGLDOKTAMSLGSPILFPAIAKELIPATEKYLGGTDDT 297
 QY 403 TEIRDSLDLGDFVFPVVALITARYHRDAGAPVYFEFRHRPOCFEDTKPAFYKADHAD 462
 DB 298 VKKDLIDLIDLADWFGVPSVIVARNHRDAGAPTYWTFQYRPSFSSDMKPKTVIGDHGD 357
 QY 463 EVRFVFGGAFKGIWMEGATEEKEKLSRKQKTKWATPARTGNPNNDLSLMPAYNLTLE 522
 DB 358 ELFSVFGAPFLK-----EGASEBEIRLSKVMKFWANFARNGNPNNGKLPHPMEYNKE 411
 QY 523 QYLQDLNMSLGGRLKEPRVDFTWS 547

DB 412 GYLQIGANTQAQKTKDKREVAFTTN 436
 Search completed: June 15, 2005, 09:43:54
 Job time : 30 secs

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 15, 2005, 09:43:26 / Search time 7228 Seconds
(without alignments)
3894.920 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079
Sequence: 1 MPGLTSSASQWCFLLQP.....PLSLTFLSLQPPFFCAP 581

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cgn2_1/USPTO.epool/6664091/runat_14062005_133640_10461/app.query.fasta.1.775
-DB=genbml -QMT=fastrcp -SUPFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091@cgn.1.14198@runat_14062005_133640_10461 -NCPU=6 -ICPU=3
-NO MMAP -LARGESOURCE -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Genbml:
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	100.0	1746	6	AR438765 Sequence
2	3079	100.0	1746	6	AX600126 Sequence
3	3079	100.0	2158	6	AR438764 Sequence
4	3079	100.0	2158	6	AX600124 Sequence

5	2915	94.7	2047	9	BC069501	BC069501 Homo sapi
6	2900	94.2	2232	6	AX477711	AX477711 Sequence
7	2881.5	93.6	2298	9	BC039073	BC039073 Homo sapi
8	2877	93.4	2092	6	AX746724	AX746724 Sequence
9	2877	93.4	2092	9	AX090927	AX090927 Homo sapi
10	2806	84.6	1896	6	BC069548	BC069548 Homo sapi
11	2606	84.6	1962	6	AX714052	AX714052 Sequence
12	2606	84.6	1962	6	AK056109	AK056109 Homo sapi
13	2588.5	84.1	1857	6	AX958392	AX958392 Sequence
14	2505	84.1	1716	6	CC733044	CC733044 Sequence
15	2250	73.1	1728	4	AB186392	AB186392 Canis fam
16	2094.5	68.0	1728	4	AB186393	AB186393 Mus muscu
17	2026	65.8	2139	4	AB045377	AB045377 Felis cat
18	1676.5	54.4	1739	4	AB060873	AB060873 Macaca fa
19	1304.5	42.4	2262	5	DIKTER	DIKTER Anas platyr
20	1285	41.7	1931	10	BC034178	BC034178 Mus muscu
21	1285	41.7	1937	10	BC034180	BC034180 Mus muscu
22	1285	41.7	1952	10	BC034191	BC034191 Mus muscu
23	1285	41.7	1961	10	BC024552	BC024552 Mus muscu
24	1285	41.7	1965	10	BC015290	BC015290 Mus muscu
25	1285	41.7	1974	10	BC031170	BC031170 Mus muscu
26	1282.5	41.7	1776	10	AB110073	AB110073 Mus muscu
27	1281	41.6	2012	10	BC055062	BC055062 Mus muscu
28	1276.5	41.5	1848	10	D50577	D50577 Mesocricetu
29	1271.5	41.3	1913	10	BC058815	BC058815 Mus muscu
30	1264.5	41.1	1927	10	AB010632	AB010632 Rattus no
31	1264.5	41.1	2257	5	BX934916	BX934916 Gallus ga
32	1263.5	41.0	2295	5	BC074230	BC074230 Xenopus l
33	1260.5	40.9	2846	10	BC027185	BC027185 Mus muscu
34	1257.5	40.8	2228	10	D50580	D50580 Rattus norv
35	1249.5	40.6	1942	10	BC015286	BC015286 Mus muscu
36	1239.5	40.3	2083	5	BC064228	BC064228 Xenopus t
37	1238	40.2	1856	10	AY034877	AY034877 Rattus no
38	1236	40.1	2249	5	BX934906	BX934906 Gallus ga
39	1233.5	40.1	1780	10	BC025812	BC025812 Mus muscu
40	1233.5	40.1	1786	10	BC024082	BC024082 Mus muscu
41	1233.5	40.1	1786	10	BC031295	BC031295 Mus muscu
42	1233.5	40.1	1788	10	BC026643	BC026643 Mus muscu
43	1233.5	40.1	1789	10	BC024491	BC024491 Mus muscu
44	1233.5	40.1	1790	10	BC024517	BC024517 Mus muscu
45	1233.5	40.1	1805	10	BC025537	BC025537 Mus muscu

ALIGNMENTS

RESULT 1
AR438765 AR438765 1746 bp DNA linear PAT 20-FEB-2004
LOCUS Sequence 3 from patent US 6664091.
DEFINITION AR438765
ACCESSION AR438765
VERSION AR438765.1 GI:42663740
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1746)
AUTHORS Curtis,R.A.J. and Silos-Santiago,I.
TITLES 53010, a human carboxylesterase family member and uses thereof
JOURNAL Patent: US 6664091-A 3 16-DEC-2003;
FEATURES
source Location/Qualifiers
1..1746
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3,086-252
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6
Length: 1746
Matches: 581
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-023-515-2 (1-581) x AR438765 (1-1746)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
 Db 1 ATGCACAGGAGACTTACTTATCATCTGCTTCCAAATGGGCTTTTCCGATTCCTCCAGCCC 60
 QY 21 LeuLeuGlnIleIleArgGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnArgAla 40
 Db 61 CTGTGGGACACAGACAGGTGGGAAAACCTGGCCCTTCTGCTAAAGGCCACAGAGAAC 120
 QY 41 ThrArgLeuGlyTrpIleGlnGlyValThrValLeuGlySerProValProVal 60
 Db 121 ACCAGGCTGGAGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGAGCCCTGGCTGTG 180
 QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAla 80
 Db 181 AACGTCTTCTCGGAATCCCTTGTGCTGCTCCCGCTGGAGATCCCTGGATTTACGAAC 240
 QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnAlaThrSerTrpProAsnLeuGly 100
 Db 241 CCGCAGCCTGCATCGCCCTGGGATTACTTGGAGAGACCACTCTACCTTAATTTGTGC 300
 QY 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnIleMetLeuLysValHisTrpProLys 120
 Db 301 CTCAGAACTCAGAGTGGCTGCTTGAATCAACACATGCTCAAGGTGATTAACCCGAAA 360
 QY 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
 Db 361 TTCGAGGTCTCAGAAAGACTGCTCTTACCTGAACATCTTAAGCCCTGCCCACGCCATACA 420
 QY 141 GlySerLysLeuProValLeuValTrpPheProGlyValAlaPheLysThrGlySerAla 160
 Db 421 GGCTCCAAAGCTCCCGCTTGTGGTGTCCAGAGAGGTGCTTCAAGACTGCTCAGCC 480
 QY 161 SerIlePheAspGlySerAlaLeuAlaIleValIleValIleValIleValIleValIle 180
 Db 481 TCCATCTTGAATGGATCCGCCCTGGCTGCTTGAAGACAGTGGTGTGTGCTCTCAG 540
 QY 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpTrpAspGlnHisAlaProGlyAsnTrp 200
 Db 541 TACCGGCTAGGAATTTGGTCTTCTTCAACATGGATCGACATGCTCCGGGGAACCTG 600
 QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGlnPhePheGly 220
 Db 601 GCCTTCAAGGACACAGTGTGCTGTCTGCTGGGTCAGAGAACTGAGATTTCTTCGT 660
 QY 221 GlyAspProSerSerValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSer 240
 Db 661 GGGGACCCCACTGTGATCCATCTTTGGCAGTCCGGGAGCCATTAAGTCTTCTAGT 720
 QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisLysValAlaIleMetGlnSerGlyVal 260
 Db 721 CTTAATACGTGTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 780
 QY 261 AlaIleIleProTrpLeuGlnAlaHisAspTrpGlnLysSerGlnAspLeuGlnValVal 280
 Db 781 GCCATCATCCCTTACCTGAGAGCCCATGATTAAGAGAGAGTGGACCTGCAAGTGGT 840
 QY 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
 Db 841 GCACATTTGTGGTAACAATGCGCTGACACTCTGAGGCCCTGAGAGTGTGCTGAGACA 900
 QY 301 LysProSerLysGlnLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVal 320
 Db 901 AAACCTCCCAAGAGGCTGACCTCAAGCCACCAAGAAACAAAGTCTTTCACCTGAGTGGT 960
 QY 321 AspGlyValaPhePheProAsnGlnProLeuAspLeuSerGlnLysAlaPheLysAla 340
 Db 961 GATGGTCTTCTTCTTAATGAGCTCTAGATTAATGCTCTCAAGAAACATTTAAAGCA 1020
 QY 341 IleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLysGln 360
 Db 1021 ATTCTTTCATCATCGAGTCAATTAACCAAGAGTGTGCTTCTGCTGCTATGAGAGAG 1080

QY 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLeuIleGlnAsnIle 380
 Db 1081 GCTCTGAGATCTCTAGTGGCTCCAAAGATCCCTTGCCCTCCATCATGATCAAAACATC 1140
 QY 381 LeuHisIleProProGlnTrpLeuHisLeuValAlaAsnGlnTrpPheHisAspLysHis 400
 Db 1141 CTGCACATCCCGCTCAGATATTGACCTTGTGGCTAATGAATATCTTCACAGAACAGC 1200
 QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlnLysAspValPhePheValVal 420
 Db 1201 TCCCTGATGAAATCCAGACAGTCTTCTGACCTTCTGAGAAATGTGTTTGTGTGCTC 1260
 QY 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGln 440
 Db 1261 CCTGACATGATCAACAGCTCGATATCAACAGATCTGGTGACCTGTCTACTTCTATGAG 1320
 QY 441 PheArgHisArgProGlnCysPheGlnAspThrLysPheValPheValIleValHis 460
 Db 1321 TTTGGCACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTGTCAAGCCGACAC 1380
 QY 461 AlaAspGlnValaArgPheValPheGlyValaPheLeuLysGlyAspIleValMetPhe 480
 Db 1381 GCTGATGAAGTCCGCTTGTGTGCTGAGTGTGCTTCTGAAAGGGGACATTTATGCTTC 1440
 QY 481 GlnGlyValaThrGlnGlnGlnGlnLysLeuLeuSerArgLysMetLeuLysTrpAlaThr 500
 Db 1441 GAAGAGCCACGGAGGAGAGAGATTACTGAGCCGGAAGATGAATGAAATCTGGGCTAAC 1500
 QY 501 PheAlaArgThrGlyAsnProAsnGlnValAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
 Db 1501 TTTGCTGAAACCGGAAATCTTAATGGAGAGACTGTCTGTGGCCACTTAATCTG 1560
 QY 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
 Db 1561 ACTGAGCAGTACCTCCAGCTGACCTTGACATGAGCTCGAGACAGAGACTCAAGAACCG 1620
 QY 541 ArgValaAspPheTrpThrSerTrpIleProLeuIleLeuSerLysSerAspMetLeuHis 560
 Db 1621 CGGGTGAATTTTGGACACACACATCCCTTATCTGCTGCTCCAGACATGCTCCAC 1680
 QY 561 SerProLeuSerSerLeuThrPheLeuSerSerLeuLeuGlnProPhePhePheCysAla 580
 Db 1681 AGTCTCTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
 QY 581 Pro 581
 Db 1741 CCT 1743

RESULT 2
 AX600126 1746 bp DNA linear PAT 14-FEB-2003
 LOCUS Sequence 3 from Patent WO0250256.
 DEFINITION AX600126
 ACCESSION AX600126.1 GI:28400192
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
 CURTIS, R.A. and SIlOs-Santiago, I.
 53010, a novel human carboxylesterase family member and uses
 thereof
 Patent: WO 0250256-A 3 27-JUN-2002;
 Millennium Pharmaceuticals, Inc. (US)
 location/Qualifiers
 1..1746
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

JOURNAL
 FEATURES
 source
 ORIGIN

Alignment Scores:

Pred. No.: 3,086-252
 Score: 3079.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 6

Length: 1746
 Matches: 581
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-10-023-515-2 (1-581) x AK600126 (1-1746)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuLLeuGlnPro 20
 DB 1 ATGCACAGGAGCTACTTCACTCTCTTCAACAATGAGTCTTTCTTGATCTCCAGCC 60
 QY 21 LeuLeuGlyHISArgGlnTrpGlyLysThrGlyProSerAlaGluGlyProGlnArgAsn 40
 DB 61 CTGTTGGGACACAGCAGTGGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
 QY 41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
 DB 121 ACCAGGCTGGAGTGGATTCAAGGACAGCAAGTCACTGCTGGAGAGCCCTGCTGTG 180
 QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
 DB 181 AACGTTCTCCGAGTCCCTTTGCTGCTCCCGCTGGAGATCCCTGATTTACGAAC 240
 QY 81 ProGlnProAlaSerProTrpAspAsnLeuArgGluAlaThrSerTrpProAsnLeuCys 100
 DB 241 CCGCAGCTGCATCCCTCCGGATTAATCTGCAAGAGCACCTCTACCTTAATTTGTC 300
 QY 101 LeuGlnAsnSerGlyTrpLeuLeuLeuAspGlnHISmetLeuLysValHISTrpProLys 120
 DB 301 CTCGAAACTCAGAGTGGCTGCTCTTAATCAACAACCTCAAGAGTGCATTACCCGAA 360
 QY 121 PheGlyValSerGlyAspCysLeuTrpLeuAsnIleTrpAlaProAlaHISAlaPheThr 140
 DB 361 TTCGAGTGTCAAGAACTGCTCTTACCTGAACATCTATGGCGCTGCCACGCCATACA 420
 QY 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
 DB 421 GGCTCCAAAGCTCCCGCTGTGGTGGTCCAGAGAGGTGCTTCAAGCTGGCTCAGCC 480
 QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlyAspValLeuValValGln 180
 DB 481 TCCATCTTTGATGGTCCGCCCTGGCTGCTATGAGAGCTGCTGTTGTGCTCCAG 540
 QY 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHISAlaProGlyAsnTrp 200
 DB 541 TACCGGCTAGGAATTTGGTTCTTCAACAATGGGATCAGCATGCTCCGGGAACTGG 600
 QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGluPhePheGly 220
 DB 601 GCCTTCAGAGACACAGTGGCTGCTGCTGCTGGGTCAGAGAAACATGCAAGTTCTTCGG 660
 QY 221 GlyAspProSerSerValThrIlePheGlyGlySerAlaGlyValAlaIleSerValSerSer 240
 DB 661 GGGGAGCCCAAGCTCTGAGCAATCTTTGGCAGTCCGGCGGACCAATAGATTTCTTCA 720
 QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHISLysAlaAlaIleMetGlySerGlyVal 260
 DB 721 CTTATACCTGCTCCCATGAGCCAAAGCTTATTCCAAAAGCATATGAGAGTGGGGTG 780
 QY 261 AlaIleIleProTrpLeuGlyAlaHISAspTrpGlyLysSerGlyLysAspLeuGlnValVal 280
 DB 781 GCCATCATCCCTTACCTGAGGCCCATATGATTAAGAGAGAGAGACCTGCAAGTGGTT 840
 QY 281 AlaHISpHeCysGlyAsnAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThr 300
 DB 841 GCACATTTCTGTGGTAACAATCGTCAAGCTTGAAGCCCTCTGAAGTGGCTTGAAGACA 900
 QY 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysTrpLysSerPheThrArgValVal 320
 DB 901 AAACCTCCAAAGAGACTGTGACCTCAAGCAGAAACAAAGCTTTTCACTCGAGTGGTT 960

QY 321 AspGlyAlaPhePheProAsnGlyProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
 DB 961 GATGATGCTTTCTTTCTTCAATGAGCTTACATCTATTTGCTCAGAAAGCATTTAAAGA 1020
 QY 341 IleProSerIleIleGlyValAsnAsnHISGlyCysGlyPheLeuLeuProMetLysGly 360
 DB 1021 ATTCCTTCATCATCGAGTCAATTAACAGAGTGGCTTCTCTGCTCATTAAGAGAG 1080
 QY 361 AlaProGlyIleLeuSerGlySerAsnLysSerLeuAlaLeuHISLeuIleGlnAsnIle 380
 DB 1081 GCTCCAGATCCCTCAAGTGGCTCCCAAGATCCCTTGGCCCTTCATGATTAACAAACATC 1140
 QY 381 LeuHISLleProProGlnTrpLeuHISLeuValAlaAsnGlyTrpPheHISAspLysHIS 400
 DB 1141 CTGCACATCCCGCTCAAGTATTTGCACTTTGCTGCTTAATGAATCTTCATGACAGACAC 1200
 QY 401 SerLeuThrGlyIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
 DB 1201 TCCCTGACTGAATCCGAGACAGTCTTGGACTTGGAGATGTGTTGTGTGTC 1260
 QY 421 ProAlaLeuIleThrAlaArgTrpHISArgAspAlaGlyAlaProValTrpPheTrpGly 440
 DB 1261 CTGCACACTGATCACAGCTCGATATCACAGAGATGCTGTGACCTGTCTACTTATAGAG 1320
 QY 441 PheArgHISArgProGlnCysPheGlyAspThrLysProAlaPheValLysAlaAspHIS 460
 DB 1321 TTTCGACCCGCTCAAGTCTTGAACACAGAGCCGCTTTTGTCAAGCCAGCACAC 1380
 QY 461 AlaAspGlyValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
 DB 1381 GCTGATGAAGTCCGCTTTGTGTGGTGTGCTTCTGAAGGGGAGCATTTGTATGTTTC 1440
 QY 481 GluGlyAlaThrGluGluGlyLeuLeuSerArgLysmetMetLysTrpAlaThr 500
 DB 1441 GAAGGAGCCACGAGAGAGAGAGTACTGAGCCGGAAGATGATAAATACGGGCTACC 1500
 QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
 DB 1501 TTGCTCAACCGGAAATCTTAATGGAAACGACTGTCTGTGGCCAGCTTATTAATCTG 1560
 QY 521 ThrGluGlnTrpLeuGlnLeuAspLeuAsnmetSerLeuGlyGlnArgLeuLysGlyPro 540
 DB 1561 ACTGAGCAGTACCTTCACCTGAGACTTGAACATGAGCTCCGACAGACTCAAGAACCG 1620
 QY 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHIS 560
 DB 1621 CGGGTGAATTTTGGACACGACCATCCCTGATCTGTCTGCTCCGACATGCTCCAC 1680
 QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePheCysAla 580
 DB 1681 AGTCTCTTTCTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
 QY 581 Pro 581
 DB 1741 CCT 1743

RESULT 3
 AR438764 2158 bp DNA linear PAT 20-FEB-2004
 LOCUS AR438764
 DEFINITION Sequence 1 from patent US 6664091.
 ACCESSION AR438764
 VERSION AR438764.1 GI:42663739
 KEYWORDS
 SOURCE
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 2158)
 AUTHORS Curtis,R.A.J. and Silos-Santiago, I.
 TITLES 53010, a human carboxylesterase family member and uses thereof
 JOURNAL Patent: US 6664091-A 1 16-DEC-2003;
 FEATURES Location/Qualifiers
 source 1..2158

ORIGIN /organism="unknown"
/mol_type="genomic DNA"

Alignment Scores:

Pred. No.:	4,08e-252	Length:	2158
Score:	3079.00	Matches:	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-023-515-2 (1-581) x AR438764 (1-2158)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleuGlnPro 20
DB ATCCACAGGAGACTTACTTATCTGCTTCACAAATGGGTCTTTTCCATTCCTCCAGCC 155
QY 21 LeuLeuGlyNH₂ArgGlnTrpGlyLysThrGlyProSerAlaGluGlyProGlnArgAsn 40
DB CTGTGGGACACAGACAGTGGGGAATACTGGCCCTTCCTGTAAGGGCCACAGAGAAC 215
QY 41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
DB ACCAGGCTGGAGATGATCAGAGGCAAGTCACTGTCTGGGAAGCCCTGTGCTGTG 275
QY 61 AsnValPheLeuGlyValProPheAlaProProLeuGlySerLeuArgPheThrAsn 80
DB AACGTGTCTCTGGAGATCCCTTTGTGTCTCCCGCTGGATCCCTGGATTTCAGAAC 335
QY 81 ProGlnProAlaSerProTrpPhePheLeuArgGluAlaThrSerLysProAsnLeuGly 100
DB CCGAGCCCTGCAATCGCCCTGGATAACTTGCAGAAAGCACTCTCACTTAATTTGAGC 395
QY 101 LeuGlnAsnSerGlyTrpLeuLeuLeuAspGlnNH₂MetLeuLysValNH₂TrpLys 120
DB CTCACAGACTCAGAGTGGCTCTTAAATCAACACATGCTCAAGTGCATTACCCGAA 455
QY 121 PheGlyValSerGlyAspGlyLeuTrpLeuAsnIleTyrAlaProAlaNH₂AlaAspThr 140
DB TTCGAGAGTCAAGAGACTGCTTACCTGAACATCTAGGCGCCCTCAGCCGATCA 515
QY 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
DB GAGCTCAAGCTCCCGCTGTGGTGTGTTCCAGAGAGGTGCTTCAGAACTGAGCTCAGCC 575
QY 161 SerIlePheAspGlySerAlaLeuAlaTyrGlyAspValLeuValValValGln 180
DB TCCATCTTTGATGGGTCCGCCCTGGCTGCTCATGAGACGTGGTGTGTGTGCTCAG 635
QY 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnNH₂AlaProGlyAsnTrp 200
DB TACGGGCTAGGAATAATTTGGTTCTTCACACATGGGATCAGCATGCTCCGGGAACTGG 695
QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpAlaGlnLysAsnIleGluPhePheGly 220
DB GCCTTCAAGACACAGGTGGCTGCTGTCTGGGTCCAGAAACATCGAGTCTTCGTG 755
QY 221 GlyAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSer 240
DB GGGGACCCCAAGCTGTGACATCTTGGCGAGTCCGCGGGAAGCATTAGTTTCTAGT 815
QY 241 IleuIleLeuSerProMetAlaLysGlyLeuPheNH₂LysAlaAlaMetGlySerGlyVal 260
DB CTTATATCTGTCTCCCAATGGCCAAAGGCTTATTCACAAACCATCAATGAGAGTGGGGTG 875
QY 261 AlaIleIleProGlyLeuGluAlaNH₂AspTrpGlyLysSerGlyAspLeuGlnVal 280
DB GCCATCATCCCTTAACCTGAGGCCCATGATTATGAGAGTGAAGCTGCAAGGTGGT 935
QY 281 AlaHisPheCysGlyLysAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThr 300
DB GCACATTTCTGTGTAAACAATGGGTGAGACTGAGGCGCTGTGAGGTGCTGAGAGACA 995

QY 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVal 320
DB AAACCTCCAGAGACTGCTGACCTCAGCCAGCAAAAGTCTTTCATCTGAGTGTCTT 1055
QY 321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysValaPheLysAla 340
DB GATGGTGTCTTCTTCTTAATGAGCTCTAGTCTATGTCTCAGAAAGCATTTAAACA 1115
QY 341 IleProSerIleIleGlyValAsnAsnNH₂GluCysGlyPheLeuLeuProMetLysGlu 360
DB ATTCCTTCATCATCTGAGTCAATTAACCAAGAGTGTGCTTCCGTGCTGCCTATGAAGAG 1175
QY 361 AlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuNH₂SteuIleGlnAsnIle 380
DB GCTCTGAGATCTTCATAGTGGTCCCAAGACGCTTGTGCTCCATCTGATACAAACATC 1235
QY 381 LeuNH₂IleProProGlnTrpLeuNH₂LeuValAlaAsnGluTrpPheNH₂AspLysNH₂ 400
DB CTGACATCCCGCTCAGTATTTGACCTTGTGGCTAATGAAATCTTCATGACAAACAC 1295
QY 401 SerLeuThrGluIleAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
DB TCCCTGACTAAATCCAGACAGCTTCTGAGCTTGTGAGATGTGTTCTTTGTGCTC 1355
QY 421 ProAlaLeuIleThrAlaArgTyrNH₂ArgAspAlaGlyValaProValTyrPheTyrGlu 440
DB CTTCCACTGATCAACAGCTCATATCAAGAGATGTGTGTGCTCACTGTCTTATGAG 1415
QY 441 PheArgHisArgProGlnCysPheGlyAspTrpLysPheProAlaPheValLysAlaAspNH₂ 460
DB TTTTGGACCGGCTCAGTCTTGAAGACACAAACCGGCTTTTGTCAAAGCCGACAC 1475
QY 461 AlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
DB GCTGATGAAGTCCGCTTGTGTGAGGTGCTTCTGAAGGAGGACATTTGATGTTTC 1535
QY 481 GluGlyAlaPheArgGluGluGlyLysLeuLeuSerArgLysMetLysTyrTrpAlaThr 500
DB GAAAGAGCCAGGAGGAGGAAATTAATGAGCCGGAATGATGAATAATCTGGCTACC 1595
QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyrAsnLeu 520
DB TTTTGTCTGAACCGGGAATCTTAATGGAAACAGCTGTCTGTGGCCACTTATATCTG 1655
QY 521 ThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
DB ACTGACAGTACTCTCAGCTGGACTTGAAACATGAGCTCGAAGAGACTCAAGAACCG 1715
QY 541 ArgValAspPheTrpPheSerThrIleProLeuNH₂LeuSerAlaSerAspMetLeuNH₂ 560
DB CGGGTGAATTTTGAACAGCACATCCCTGAACTGTGTGCTCGCTCGACATGCTCCAC 1775
QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
DB AGTCCCTTTCTTCTTAATCTTCTCTCTCTCTCTCCAGCCTTTCTTTCTTTGTGCT 1835
QY 581 Pro 581
DB 1836 CCT 1838

RESULT 4
AX600124 2158 bp DNA linear PAT 14-FEB-2003
LOCUS Sequence 1 from Patent WO0250256.
DEFINITION
ACCESSION AX600124
VERSION AX600124.1 GI:28400190
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1

AUTHORS Curtis, R.A. and Silos-Santiago, I.
TITLE 53010, a novel human carboxylesterase family member and uses
thereof
JOURNAL Patent: WO 0250256-A 1 27-JUN-2002;
Millennium Pharmaceuticals, Inc. (US)

FEATURES
source
1. 2158
Location/Qualifiers

CDS

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
96.1841
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD66643.1"
/db_xref="GI:28400191"
/translation="MPQSLTSSASQWCFLLIQPLGHRQKCTPSAEGPQRTRLG
MIQKQVTVLSPVFNVPFLGVFPAPAPGSLRFTNPSPASBPMDLRKATSPNLCQ
NSEWLLDQHLKVPKPGVSEDLNITYAPAHADQSDKLPVLVWPPGAFKXGSA
SIPDSSALAAEDVLVVVQYRLGIFGFTWDDQAPGNMVAFKDQVVALSWQKIEF
FGDPSSTYIFGESAGAISVSLSILSPAKGLFKHAIMESGVAILPYLEADHYKSED
LOVAFPGNNADEBALRCLRTKPSKELLTLQKTSFPRVNDGAPPNPDLIS
QKARAIPTSIGVNNHCEGFLPMKEAPBILSGSKSLALHLIQNTLHPLOYHLVA
NEYFDKSLTEIRDSLDDLGDFVFPALITARYHADAGAPVFFYFRHRPQCFED
TKPAVKADHADVPAFVGAFPLKGDIVMFBGATEERKLSKMKMYATAPARTGNPN
GNDLSIMPAYNLTBOYLQDLNMSLGQRLEKPRVDFMTSTIPLILASDMLHSPLSL
TFLSLQPFPPFCAP"

ORIGIN

Alignment Scores:

Pred. No.: 4,086-252 Length: 2158
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-023-515-2 (1-581) x AX600124 (1-2158)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTTPCySPhePheLeuIleLeuGlnPro 20
DB 96 ATGCCACAGGAGCTTACATTCATCTCTCACAAATGATGCTTTTCTGATTTCTCAGCC 155
QY 21 LeuLeuGlnIleSarGlnTTPQIlySerThrGlyProSerAlaGlnGlyProGlnArgAn 40
DB 156 CTGTTGGGACACAGCAGTGGGGGAAACTGGGCTTCTCGTCGAGAGGCCACAGAGAAC 215
QY 41 ThrArgLeuGlyTTPIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
DB 216 ACCAGGCTGGGATGATTCAGGGCAAGCAAGTCACTGATCTGGGAAGCCCTGTGCTGTG 275
QY 61 AsnValPheLeuGlyValProPheAlaAlaProPoleuGlySerLeuArgPheThrAn 80
DB 276 AACGTGTCTCCGAGTCCCTCTTGTCTGCCCGCTGGGATCCCTGGGATTTACGAAC 335
QY 81 ProGlnProAlaSerProTTPAspAsnLeuArgGlnAlaThrSerTyrProAsnLeuCyS 100
DB 336 CGCAGCCTGCATCGCCTGGGATTACTTGCAGAAAGCCACTCTACCTTAATTGTGTC 395
QY 101 LeuGlnAsnSerGlnTTPLeuLeuLeuAspGlnIleMetLeuLysValIleTyrProLys 120
DB 396 CTCGAGAACTCAGAGTGGCTGCTTAAATCAACACATGCTCAAGTGCATTAACCGAAA 455
QY 121 PheGlyValSerGlnAspCysLeuTyrLeuAsnIleTyrAlaProAlaIleAlaAspThr 140
DB 456 TTCGAGGTCTCAGAACTGCTCTTAACCTGAACATCTATGCGCTGCCACGCCCATACA 515
QY 141 GlysSerLysLeuProValLeuValTrrPheProGlyGlyAlaPheLysThrGlySerAla 160
DB 516 GAGCTCAAGCTCCCGCTCTGGTGTGTGTTCCAGAGAGGTGCTTCAAGCTGGCTCAGCC 575
QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTyrGlnAspValLeuValValValGln 180
DB 576 TCCATCTTGAATGGGTCCGCGCTGGCTGCTATGAGGAGCGTCTGTGTGTGCTGTCCAG 635

QY 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTTPAspGlnIleAlaProGlyAsnTTP 200
DB 636 TACCGGCTAGGAATATTGTTGTTCTTCAACACATGGGATCAGCAAGCTCCGGGAACTGC 695
QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTTPValGlnLysAsnIleGlnPheGly 220
DB 696 GCCTTCAAGAGACAGAGTGGCTGCTCTGCTGGGTCAGAAAGAACATGAGATTCTTCGCT 755
QY 221 GlysAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSer 240
DB 756 GGGAGCCCGACCTGTATCAATCTTTGGCAGTCCGGGAGCCATTAAGTGTTCATGCT 815
QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheIleLysAlaIleMetGlySerGlyVal 260
DB 816 CTTAATCTGTCTCCCATAGGCCCAAGGCTTATTCCAAAGCCATATGAGAGTGGGGT 875
QY 261 AlaIleIleProTyrLeuGlnValAlaIleAspTyrGlnLysSerGlnAspLeuGlnVal 280
DB 876 GCCATCATCCCTTACCTGAGGCCCATGATTATGAGAAAGATGAGACCTGCAAGGTGTT 935
QY 281 AlaIlePheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
DB 936 GCACATTTCTGTGTGTAACAATGCGTCAGACATCTGAGGCCCTGTGAGTGCCTGAGACA 995
QY 301 LysPProSerLysGlnLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVal 320
DB 996 AACCTCTCAAGAGACTGCTGACCTTCAGCCAGAAACAAAGCTTTTCACTCGAATGGTT 1055
QY 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuSerGlnLysAlaPheLysAla 340
DB 1056 GATGTGCTTTCTTCTTCCAAAGAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGA 1115
QY 341 IleProSerIleIleGlyValAsnAsnIleGlnCysGlyPheLeuLeuProMetLysGln 360
DB 1116 ATTCCTTCATCATCGAGATCAATTAACAGAGTGTGCTTCTGCTCCTCATATGAGAG 1175
QY 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuIleGlnAsnIle 380
DB 1176 GCTCTGAGATCTCAGTGGCTCCCAACAAAGTCCCTTGCTTCATCTATACAAACATTC 1235
QY 381 LeuIleIleProProGlnTyrLeuIleLysValAlaAsnGlnTyrPheIleAspLysIle 400
DB 1236 CTGACATCCCGGCTCAGTATTGCACTTGAGGTATGATATCTTCATGACAGAGCAC 1295
QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
DB 1296 TCCCTGACGTGAATCCGAGACAGTCTTCTGACCTTGAGAGATGCTTCTTGTGTGTC 1355
QY 421 ProAlaLeuIleThrAlaArgTyrIleSarAspAlaGlnValAlaProValTyrPheTyrGln 440
DB 1356 CCTGACATGATCAACGCTCGATATCAAGAGATGCTGTGACCTGTCTACTTCTATAGAG 1415
QY 441 PheArgIleSarProGlnCysPheGlyAspThrLysProAlaPheValIleAspIle 460
DB 1416 TTTCGGCACCGGCTCAGTCTTGAAGACACAGAAAGCGGCTTTGTCTAAAGCCGACAC 1475
QY 461 AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
DB 1476 GCTGATGAAGTCCGCTTGTGTGCTGTGCTGCTTCTGAAAGGGGAGCATTTGTTGTTTC 1535
QY 481 GlnGlyValaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTyrTrrAlaThr 500
DB 1536 GAAGGAGCCAGAGAGAGAGAGATTAATGAGCCGAGAAAGATGAATAATACGGGCTACCC 1595
QY 501 PheAlaArgThrGlyAspProAsnGlyAsnAspLeuSerLeuTTPProAlaTyrAsnLeu 520
DB 1596 TTTCGTCGAACCGGGAATCTTAATGGAGACGCTGTCTCTGTGCCAGCTTATATGTC 1655
QY 521 ThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
DB 1656 ACTGAGAGTACCTTCAGCTGAGCTTGAACATGAGCCTCGGACAGAGACTCAAGAAACCG 1715

	FEATURES
Qy	541 ArgValAspPheThrPhrSerThrIleProLeuLlleLeuSerAlaSerAspMetLeuHis 560
Dd	1716 CGGGTGGATTTTGGACGAGCACCACTCCCGTCGTCTTGCTGCCTCCACACTGCTCCAC 17175
Qy	561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
Dd	1776 AGGCTCTTCTTCTCTTAATTCTTCCCTCTCTCTCCACGCCCTTTCTTTTCTTTTGCT 18355
Qy	581 Pro 581
Dd	1836 CCT 1838
RESULT 5	
LOCUS	BC069501 2047 bp mRNA linear PRI 19-AUG-2004
DEFINITION	Homo sapiens hypothetical protein FLJ31547, mRNA (CDNA clone
ACCESSION	MGC:97182 IMAGE:7262427), complete cds.
VERSION	BC069501
KEYWORDS	BC069501.1 GI:46854682
SOURCE	MGC.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2047) Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altchul,S.F., Zeeberg,B., Buetow,K.H., Scheefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Ditchenko,L., Marusik,K., Farmer,A.A., Rubin,G.M., Hong,L., Stepienon,M., Soares,W.B., Bonaldo,M.F., Casavant,T.L., Schaeff,T.E., Brownstein,M.J., Udell,T.B., Tsohyiki,S., Carminci,P., Prange,C., Raha,S.S., Loquellano,N.O., Peters,G.J., Abrahamson,R.D., Mullahy,S.J., Bosak,S.A., McGowan,P.J., McMernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.C., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fatey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouford,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schwartz,T., Myers,R.M., Bretfield,Y.S., Krzywinski,M.I., Skalski,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.U. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932
TITLE	2 (bases 1 to 2047) Director MGC Project.
JOURNAL	Direct Submission
PUBMED	Submitted (29-AR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
AUTHORS	NIH-MGC Project URL: http://mgc.nci.nih.gov
TITLE	Contact: MGC help desk
JOURNAL	Email: gcgaps-remail.nih.gov
REMARK	Tissue Procurement: Baylor Human Genome Sequencing Center CDNA Library Preparation: Baylor Human Genome Sequencing Center CDNA Library Arrayed by: The I.M.A.G.E. Consortium (NLN) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne,P.H., Garcia,A.M., Lu,X., Hulyk,S.W., Louieged, H., Kowal,C.B., Sneed,A.J., Martin,R.G., Muzny,D.M., Nanavati, A.N., Gibbs,R.A.
COMMENT	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Series: IRRA Plate: 4 Row: e Column: 1.
FEATURES	Location/Qualifiers

[illegible]

172 GluapervallleuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr 191
 616 GAGGAGCGTCTGTTGTGTGCTCCAGTACCGGCTAGGAATTTGGTTCTTCCACACA 675
 192 TTPAEPGLNHIALPProGlyAsnTPRALaPheLyAspGlnValAlaAlaLeuSerTrp 211
 676 TGGGATCAGACGCTCCGGGGAACGTGGGCTTCAAGACACAGGTGGCTGCTGTCCTGG 735
 212 ValGlnLyAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGln 231
 736 GTCCGAAAGAACATGATGATTTCTTCGTTGGGAGCCCACTCTGTATCCATTTTGGCGAG 795
 232 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLyGlyLeuPhe 251
 796 TCCGCGGAGACCATAGTATTTCTAGTCTTATCTGCTCTCCCATGGCCAAAGCTTATTCC 855
 252 HisLyAsnAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyr 271
 856 CACAAAGCCATCATGAGAGTGGGGTGGCCATCTCCCTTACCTGAGGCGCATGATTAT 915
 272 GluLySerSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 291
 916 GAGAAAGATGAGAGACCTCGAGGTGGTTGCACTTCTGTGTAAACATGCTCAAGACTCT 975
 292 GluAlaLeuLeuArgCysLeuArgThrLyPProSerLyGlnLeuLeuThrLeuSerGln 311
 976 GAGGCCCTCTCTAGGTGCTCGAGACAAACCTCCAAAGACCTGCTGACCTCCAGCCAG 1035
 312 LyThrLySerPheThrThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 331
 1036 AAAACAAAGTCTTTCATCTCGATGAGTGGTGGTCTTCTTCTTAATAGCCCTCAGAT 1095
 332 LeuLeuSerGlnLyAlaPheLyAsnAlaIleProSerIleIleGlyValAsnAsnHisGln 351
 1096 CTATTGTCTCAAAAGCATTTTAAAGCAATTTCTTCCATCATGAGTCAATTAACACAGAG 1155
 352 CysGlyPheLeuLeuProMetLySerGluAlaProGlnIleLeuSerGlySerAsnLySer 371
 1156 TGTGCTTCTCTGCTGCTCTGATGAGAGGCTCCTGAGATCTCGAGTCCCAACAGTCC 1215
 372 LeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIleLeuHisLeuVal 391
 1216 CTTCGCCCTCATCTGATACAAACATCTCGACATCCGCGCTCAGATATTGACCTTGG 1275
 392 AlaAsnGlnTyrPheHisAspLySerHisSerLeuThrGlnIleArgAspSerLeuLeuAsp 411
 1276 GCTAATGATATCTTCCATGACAAAGCATCTCCGACTGAATCCGAGACAGTCTTCTGAG 1335
 412 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAsp 431
 1336 TTGCTTGAGATGTGTTCTTGTGTGCTCCGTGACATGATCAGCTGATATCACAGAGAT 1395
 432 AlaGlyAlaProValTyrPheTyrGlnPheArgHisArgProGlnIleCysPheGlnAspThr 451
 1396 GTGTGTGACACCTGTCTACTTCTATGAGTTTGGGACCGGCTCAAGTCTTGAAGACAG 1455
 452 LyPProAlaPheValLyAsnHisAspHisAlaAspGlnValArgPheValPheGlyGlyAla 471
 1456 AAGCGGCTTTTGTCAAGCCACACACGCTGATGAGTCCGCTTGTGTGTGCTGGTGGC 1515
 472 PheLeuLyGlyAspIleValMetPheGlnGlyAlaThrGlnGlnGlnLyAsnLeuSer 491
 1516 TTCTGAAAGGGGACATTTTATTTATGTTCCAAAGAGCCACGAGAGAGAGAGATTACTGAGC 1575
 492 ArgLySerMetLeuLyTyrIleAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 511
 1576 CGGAGATGATGAATACTGGGCTACTTGTCTCGAAACGGGAAATCTTAATGGGAACGAGC 1635
 512 LeuSerLeuTrpProAlaTyrAsnLeuThrGlnIleTyrLeuGlnLeuAspLeuMet 531
 1636 CTGTCTCTGTGGCCAGCTTATATGACTGAGCAGATACCTCCAGCTGGAATTTGAACATG 1695
 532 SerLeuGlyGlnArgLeuLyGlnProArgValAspPheTrpThrSerThrIleProLeu 551

1696 AGCCTCGACAGAGACTCAAAAGAACCGGGGTGATTTTGGACCAAGCATCCCGCTG 1755
 552 IleLeuSerAlaAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 571
 1756 ATCTGTCTGCTCTCGAAGATGCTCCACAGTCTCTTCTTCTTAATCTTCTCTCTC 1815
 572 LeuGlnProPhePhePhePheCysAlaPro 581
 1816 CTCACGCTTCTTCTTCTTCTTGTGCTCT 1845
 RESULT 6
 AX477711 2232 bp DNA linear PAT 12-AUG-2002
 LOCUS AX477711
 DEFINITION Sequence 23 from Patent WO0246426.
 ACCESSION AX477711
 VERSION AX477711.1 GI:22216866
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS
 1 Sanjanwala,M.M., Yao,M.G., Au-Young,J., Baughn,M.R., Arvizu,C.,
 Ring,H.Z., Lee,B.A., Ding,L., Hafalia,A.J., Tang,Y.T., Yue,H.,
 Tribouley,C.M., Lu,D.A., Lal,P.G., Warren,B.A., Yang,J.,
 Wajia,N.K., Nguyen,D.B., Gandhi,A.R. and Ison,C.H.
 TITLE
 JOURNAL
 Drug metabolizing enzymes
 Patent: WO 0246426-A 23 13-JUN-2002;
 Incyte Genomics, Inc. (US)
 FEATURES
 source
 1..2232
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 6538080CB1"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,39e-237 Length: 2232
 Score: 2900.00 Matches: 548
 Percent Similarity: 99.82% Conservative: 2
 Best Local Similarity: 99.46% Mismatches: 1
 Query Match: 94.19% Indels: 0
 DB: 6 Gaps: 0
 US-10-023-515-2 (1-581) x AX477711 (1-2232)
 31 GTPProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLyGln 50
 274 GGGCTTCTCTGAAAGGGCCACAGAGAACACACAGGCTGGATGATTCAGGGCAAGCA 333
 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 70
 334 GTCACTGTGCTGGAAGGCTGTGCTGTGAACGATTCCTCGAGTCCCTTCTGTGCT 393
 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeu 90
 394 CCCCCGGGAGATCCCTCGATTAAGAACCCGACGCTGCACTGCCCTGGATTAACCTTG 453
 91 ArgGlyAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGlyTrpLeuLeuAsp 110
 454 CGAGAGCCACTCTTACCTTAATTTGTGCTCCAGAACTCAGAGACTGAGTCTCTTAAGT 513
 111 GlnHisMetLeuLyValHisTyrProLyPheGlyValSerGluAspCysLeuTyrLeu 130
 514 CAACACATGCTCAAGTCAATTACCCGAAATTCGAGAGTCAAGAGACTGCTCTTAAGT 573
 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLyLeuProValLeuValTrpPhe 150
 574 AACATCTATGGCCCTGCCACGCGCATACAGGCTCCAAAGCTCCCGTCTTGTGTGGTTC 633
 151 ProGlyGlyAlaPheLyThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 170

Db	634	CSAGAGAGTGGCTTCAAGACTGGCTCAAGCTTCCATCTTTGATGGGTCCGCCCTGGCTGCC	693
Qy	171	TryGIuAaPValLeuValValValGIuNTrArgLeuGIyIlePheGIyPhePheThr	190
Db	694	TATAGAGACGGTGGTGGTGGTGGTCCAGACCGGCTAGAAATATTGGTCTTCAC	753
Qy	191	ThrTrpAaPGrInHValAProGILyAaSTrPAlaPheIyAaPGrInValAlaAlaLeuSer	210
Db	754	ACATGGAGATCAGCATGCTCCGGGAACTGGAGCTTCAAGAACAGGGGTGGTGGTCC	813
Qy	211	TrpValGIuIyAaenIIeGIuPhePheGIyGIyAaPProSerSerValThrIlePheGIy	230
Db	814	TGGTTCAGAAAGAACATCGAGTCTTGGTGGGGACCCGAGCTGTGACCATCTTTGGC	873
Qy	231	GIuSerPAlaGIyAlaIleSerValIleSerLeuIIeIuSerProMetAlaIyGIyLeu	250
Db	874	GAGTCGGCGGAGACCATAGTGTCTTACGCTTATATGCTCTCCATGGCGAAAGGCTTA	933
Qy	251	PheHISlyAaIaIleMetGIuSerGIyValAlaIleIleProTrpLeuGIuAlaHISaP	270
Db	934	TTCCAGAAACCATCATAGAGAGTGGGGTGGCCATATCCCTTACCTGGAGGCCATGAT	993
Qy	271	TryGIuIyAaSerGIuAaPLeuGIuValAlaHISPheCysGIyAaAaanaIaSerAaP	290
Db	994	TATGAAAGAGTAGAGACCTGCGAGGGTGTTCACATTTCTGTGTAAACAATCGTCAGAC	1055
Qy	291	SerGIuAlaLeuLeuArgCysIleAaGTrHlyPProSerIyGIuLeuLeuThrLeuSer	310
Db	1054	TCGAGGCGCTGGTAGGTCCCGAGAGCAAAACCTCCAAAGAGCTGGTGAACCTTGAC	1111
Qy	311	GIuIyEThrIyAaSerPheThrArgValValAspGIyAlaPhePheProAaNGIuProIeu	330
Db	1114	CAGAAACAAAGCTTTCACCTCGAGTGGTGAATGGGTCTTCTTCCCTAATGAGCTCTA	1172
Qy	331	AaPLeuLeuSerGIuIyAaIaPheIyValIleProSerIleIleGIyAlaAaenHIS	350
Db	1174	GATCTATGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATACACAC	1233
Qy	351	GIuCysGIyPheIuLeuPProMetIyGIuAlaProGIuIleIuSerGIySerAaIyS	370
Db	1234	GAGTGGTGGCTTCTGCTGCTCATGAAGAGGCTCTCGAGTGGCTCAGTGGCTCAACAG	1292
Qy	371	SerLeuAlaIeHISIleuIleGIuAaenIIeIuHISIleIleProProGIuTrIleuHISIleu	390
Db	1294	TCCCTTGGCTCCATGTATACAAACATCTCGACATCCGCTCAGATTTTGACCTT	1353
Qy	391	ValAlaAaenGIyTrPheHISaPTrpHISerLeuThrGIuIleArgAaSerLeuLeu	410
Db	1354	GTGGCTAATGATATCTTCATGACAGACCTCCGACGTGAATCCGAGACAGCTTTCTG	1413
Qy	411	AaPLeuLeuGIyAaPValPhePheValValProAlaIleuIIeThrAlaArgTrpHISaP	430
Db	1414	GACTGTGTGGAGATGTCTTTGTGGTCCCTGACATGATCACAGCTCGATATCACAGA	1473
Qy	431	AaPAlaGIyAlaProValTrpPheTrpGIuPheArgHISaPProGIuIyAaPheGIuAaP	450
Db	1474	GATGTGGGACACCTGTACTCTTATAGATTCCGACACCGGCTCACTGCTTTGAAGAC	1533
Qy	451	ThrIySProAlaPheValIyAaPHisAlaAaPGrIuValaArgPheValPheGIyGIy	470
Db	1534	ACGAAGCGGCTTTGTCAAGCCACACGCTGATGAAGTCCGCTTGTGTGGTGGT	1593
Qy	471	AlaPheLeuIySGILyAaPILeValMetPheGIuGIyAlaThrGIuGIuGIuIyLeuLeu	490
Db	1594	GCTTCCCTCGAAGGGGACATTTGTTATGTTCAGAGGACCAAGAGAGAAAGTACTG	1653
Qy	491	SerArgIyMetMetIyETrpAlaThrPheAlaArgThrGIyAaenProAaenGIyAaen	510
Db	1654	AGCCGGAAGATGAATATCTGGGCTACTTTGTCTGAACCGGGAATCCTAATGGGAC	1713
Qy	511	AaPLeuSerLeuTrpProAlaTrpAaenThrGIuGIuTrpLeuGIuIleAaPLeuAaen	530

Dd	1714	GACCTGTTTCGTGGCCAGACTTATATCATGACTGCAGTACCCTCAGCTGGATTGAAC	1772
Oy	531	MctSerLeuGLYAlaArgLeuIysGIuProArgValAspPheTrpHisSerThrIlePro	550
Dd	1774	ATGAGCCTCGGACAGAGACTCAAAAGAACCGCGGTGGAGTTTGACCAAGACCAATCCC	1833
Oy	551	LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer	570
Dd	1834	CTGATCTCGTGTGCTCCGACATGCTCCAGAGTCTCTTCTTCTTAATCTTCTCTCT	1893
Oy	571	LeuLeuGIuProPhePhePhePheCysAlaIlePro	581
Dd	1894	CTCTCCAGCCTTCTTCTTCTTCTTGTGCTCCT	1926
RESULT 7			
LOCUS	BC039073	2298 bp	mRNA linear PRI 18-NOV-2002
DEFINITION	Homo sapiens, similar to carboxylesterase, clone IMAGE:5174937,		
ACCESSION	BC039073		
VERSION	BC039073.1		
KEYWORDS	GI:25058132		
SOURCE	.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 2298)		
JOURNAL	Strausberg,R. Direct Submission Submitted (01-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA		
REMARK	NIH-MGC project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: gcgaps@remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amgbcm.tmc.edu Guarnatine, P.H., Garcia, A.M., Lu, X., Huilyk, S.W., Loulsesged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.		
FEATURES			
SOURCE			
ORIGIN			
Alignment Scores:			
Pred. No.:	2,89e-235	Length:	2298
Score:	2881.50	Matches:	555
Percent Similarity:	90.85	Conservative:	1
Best Local Similarity:	90.69	Mismatches:	7
Query Match:	93.59%	Indels:	49
GB:	9	Gaps:	3
Location/Qualifiers			
1..2298			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:5174937"			
/tissue_type="Brain, Lung, Testis, adult, pooled whole"			
/clone_id="NH MGC_115"			
/lab_host="DH10B"			
/note="Vector: pCMV-SPORT6"			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Series: IRAP Plate: 51 Row: k Column: 18 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.			

FEATURES		biotechnology (UP)	
source		Location/Qualifiers	
		1..2092	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
ORIGIN			
Alignment Scores:			
Pred. No.:	6.16e-235	Length:	2092
Score:	2877.00	Matches:	551
Percent Similarity:	98.40%	Conservative:	2
Best Local Similarity:	98.04%	Mismatches:	2
Query Match:	93.44%	Indels:	2
DB:	6	Gaps:	2
US-10-023-515-2 (1-581) x AX746724 (1-2092)			
QY	24	HisArgGlnTrpGlySerThrGlyProSerAlaGluGlyProGlnArgAsnThrArgLeu	43
DB	126	CACTTTAACTGG-----GGGCGCTTCGCTGAGGGCCACAGAGAACCGAGCTG	176
QY	44	GlyTrpIleGlnGlyLeuGlnValThrValLeuGlySerProValProValaenValPhe	63
DB	177	GGATGGATTCAAGGCAAGCAAGTCACTGTGCTGGAAAGCCCTGTGCTGTGAAGTGTTC	236
QY	64	LeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsnProGlnPro	83
DB	237	CTCGGAGTCCCTTCCTGCTGCTCCCGCTGGATTCCTGGATTACGAACCGCAGCCT	296
QY	84	AlaSerProTrpAspAsnLeuArgGluAlaThrSerTyProAsnLeu-----Cy	100
DB	297	GCATGCGCTCGGAGTAACTGGCAGAGAGCCACTCTCACTTAATTGTAAACCAAGTG	356
QY	100	AlenGlnAsnSerGluTrpLeuLeuAsnArgIleHisMetLeuValHisTyProTy	120
DB	357	CCTCAGAACTCAGAGTGGCTGCTCTTAATCAACATGCTCAAGTGCAATTACCGAA	416
QY	120	AspGlyValSerGluAspCysLeuTyTrpLeuAsnIleTyAlaProAlaHisAlaAspTh	140
DB	417	ATTCGGAATGTACAGAAAGCTGCTCACTGAACATCTATGGGCTGGCCAGCGCATAC	476
QY	140	RGlySerTyLeuProValLeuValTrpPheProGlyGlyAlaPheLeuTyThrGlySerAl	160
DB	477	AGGCTCCAGGCTCCCGTCTTGTTGTGTGTTCACAGAGGTGGCTTCMAACATGGGCTCAAC	536
QY	160	AsenIlePheAspGlySerAlaLeuAlaAlaTyTrpGluAspValLeuValValAlaG	180
DB	537	CTCATCTTTGATGGGTCGCGCTGCGCTATAGAGACGTGTGTGTGTGTGTGTGTCA	596
QY	180	nTyArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTr	200
DB	597	GTACCGGCTAGGAATATTGTGTTCTTCAACCAATGGGATCAGCATGCTCCGGGGAACAG	656
QY	200	PalAspPheLeuAspGlnValAlaAlaLeuSerTrpValGlnTyAsnIleGluPhePheG	220
DB	657	GGCTTCAAGGACAGAGTGGCTGCTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	716
QY	220	YgLYAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSe	240
DB	717	TGGGAGACCCCAAGCTCTGTGACCATCTTGGGCAAGTCCCGGAGGACCAATAGTGTCTAG	776
QY	240	rLeuIleLeuSerProMetAlaTyGlyLeuPheHisAlaValAlaIleMetGluSerGlyVa	260
DB	777	TCTTATATCTGTCTCCATGCGCCAAAGCTTATTCCCAAAAGCATCATAGAGAGTGGGT	836
QY	260	AlaIleIleProTyTrpLeuGluAlaHisAspTyTrpGlySerGluAspLeuGlnValVa	280
DB	837	GGGCATCATCTTCACTTACGAGGCGCCATGATTATGAGAAAGTAGAGCACTGCAAGTGT	896
QY	280	AlaHisPheCysGlyAsnAsnAlaSerAspSerGluAlaLeuLeuAspGlyAlenGlyTh	300
DB	897	TGCACATTTCTGTGTAAACATGTGTCAGACTTGAAGGCTGTGTGAGAGTCCGTGAAGAC	956

[illegible]

REFERENCE
AUTHORS
1
Oe,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Makamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahashi,K., Murakami,K., Yasuda,T., Iwayanagi,T., Watanabe,M.,
Shiratori,A., Sudo,H., Hosokiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hirooka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Matsuda,S.,
Yosida,M., Hotta,T., Kusano,J., Kanohori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiyama,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Mueshino,K., Yuuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kunagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tachiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
Kajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togaishi,T.,
Oyama,M., Hara,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Maehiro,Y., Yamashita,R.,
Nagai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039
JOURNAL
PUBMED
REFERENCE
AUTHORS
2
Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Ota,T., Makamatsu,A., Ishii,S., Yamamoto,J.,
Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
Kanda,K., Watanabe,M., Murakami,K., Kanohori,K.,
Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahashi,K.,
Maehiro,Y., Nagai,K. and Isogai,T.
NEBO human cDNA sequencing project
Unpublished
3 (bases 1 to 2092)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomicehri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan, cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1..2092
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAMH2002191"
/issue_type="Brain"
/clone_lib="BRAMH2"
/note="Cloning vector: pMER18PL3"
394..1803
/note="unannotated protein product"
/codon_start=1
/protein_id="BAC03565.1"
/db_xref="GI:21749269"
/translation="MLKVHPKFGVSEDDLYLTIYAPAHADTGSRLPVLVWPPGAGFK"

ORIGIN
Alignment Scores:
Pred. No.: 6,16e-235 Length: 2092
Score: 2877.00 Matches: 551
Percent Similarity: 98.40% Conservative: 2
Best Local Similarity: 98.04% Mismatches: 2
Query Match: 93.44% Indels: 7
D8: 9 Gaps: 2
US-10-023-515-2 (1-581) x AK090997 (1-2092)
QY 24 HisArgGlnTrpGlyLeuThrGlyProSerAlaGluGlyProGlnArgAsnThrArgLeu 43
DB 126 CACTTAACTGG-----GGGCTTCTCTGAGAGGCCACAGAGAACCGAGGCTG 176
QY 44 GlyTrpIleGlnGlyLeuValThrValIleuGlySerProValProValAsnValPhe 63
DB 177 GGATGGATTCAAGGGAACCAAGTCACTGCTGGGAACCCCTGTCCTGTAAGCTGTTTC 236
QY 64 LeuGlyValProPheAlaIleProProLeuGlySerLeuArgPheThrAsnProGlnPro 83
DB 237 CTGGAGTCCCTTGTGCTGCTCCCGCGGAGATCCCGCATTTTACCAACCCGCGACT 296
QY 84 AlaSerProTrpAspAsnLeuArgGlnAlaThrSerTrpProAsnLeu-----C 100
DB 297 GCATCGCCCTGGAGTAACCTTGCAGAGACCACTCTTACCTTAATTGTAAGACCAAGTG 356
QY 100 sLeuGlnAsnSerGluTrpLeuLeuAspGlnHisMetLeuValHisGlyProGly 120
DB 357 CCTCGAAGACTCAGAGTGGCTGCTTTCATGATCAACATGCTCAAGTGATTAACCGAA 416
QY 120 sPheGlyValSerGluAspGlyLeuTrpLeuAsnIleTrpAlaProAlaHisGlnAspThr 140
DB 417 ATTCGAGATGCAGAGACTGCTCTTACCTGAACATCATATGCGCTGCCAGCGCATAC 476
QY 140 rGlySerTrpLeuProValIleuValIleuProGlyGlyValAlaPheValSerAl 160
DB 477 AGGCTCCAGAGTCCCGCTTGTGTGTGTCCAGAGAGTGCCTTCAAGACTGCGCTCAGC 536
QY 160 aSerIlePheAspGlySerAlaLeuAlaIleTrpGlyAspValIleuValAlaValG 180
DB 537 CTTCATCTTTATGGGTCCGCTGCGCTGCTTATGAGAACGTCGTGTTGCTGCTCA 596
QY 180 nTyArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHisAlaProGlyAsnTr 200
DB 597 GTACCGGTTAGGAATATTGGTTCTTCAACATGAGATGATGATGATGATGATGATGATG 656
QY 200 pAlaPheValAspGlnValAlaAlaLeuSerTrpValGlnValAsnIleGluPhePheG 220
DB 657 GGCTTCAGAGCAAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
QY 220 YGlyAspProSerSerValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSe 240
DB 717 TGGGAGCCCAAGCTCTGTGACCATCTTTGGCGAGTCCGCGGAGCGCAATAGTTTCTAG 776
QY 240 rLeuIleLeuSerProMetAlaValGlyLeuPheHisValAlaIleMetGluSerGlyVa 260
DB 777 TCTTATACTGCTCTCCATGAGCCCAAGGCTTATTCACCAAAACCATCATGAGAGAGTGGGT 836
QY 260 lAlaIleIleProTrpLeuGlnAlaHisAspTrpGlnValSerGlnValAspLeuGlnValVa 280
DB 837 GGCCATCATCCCTTAACCTGAGAGCCCAAGATTAAGAAAGTAGAGAGCTGCAAGTGTGT 896
QY 280 lAlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgTh 300

```

Db      897  TGCACATTTCTGCTGTAACAATGTCAGACTGAGCCCTGCTAGAGCTGCTGAGAC
Qy      300  rlysrProserLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVa
Db      957  AAAACCTCCCAAGAGCTGCTGACCCCTCAGCCAGAAAACAAAGCTTTCCTGCTGAGTGGT
Qy      320  laeGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAl
Db      1017  TGATGGTGGCTTTCTTCTTAATAGCTTCAAGCTTCAATATCTATCTCTCAGAAAGCATTTAAAGC
Qy      340  alieProserLileGlyValAsnAsnHisGluCysGlyPheLeuLeuProMetLysG1
Db      1077  AATTCCTTCATCATGAGGTCAATAGCCACGAGTGGCTTCTCTGCTGCTGCTATGAAAGA
Qy      360  uAlaProGluLeuLeuSerGlySerSerMetLysSerLeuAlaLeuHisLeuLileGlnMet1
Db      1137  GGCTCTTGAGATCTCTAGTGGCTCCCAACAAAGTCCCTTGCCCTTCATGTATCAAAACAT
Qy      380  eleuHisLileProProGlnLysLeuHisLeuValAlaAsnGluLysPheHisAspLysHi
Db      1197  CCTGCACATCCCGCCCTCAATATTGCACTTGCTGCTAATGAACTTCCATGACAGACCA
Qy      400  sSerLeuThrGluLileArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVa
Db      1257  CTCCCTGACTGAATCCGAGACAGCTTCTGAGCTTGAGAGATGTGTCTTTGTGGT
Qy      420  lProAlaLeuLileThrAlaArgTyPHisArgAspAlaGlyAlaProValLysPheTyG1
Db      1317  CCTGCACATGATCACAGCTCGATATCACAGAGATGCTGAGACCTGCTCTACTTTATGA
Qy      440  uPheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAspHi
Db      1377  GTTTCGGGACCCGCTCAGTGTCTTGAAGACACGAAAGCCGGCTTTGTCAAGCCGACCA
Qy      460  aAlaAspGlyValaLysPheValPheGlyGlyAlaPheLeuLysGlyAspLileValMetPn
Db      1437  CCTCATATGAATCCGCTTGTGTGTGGTGGTGGTCTTCTGAAAGGGGACATTTGATATCTT
Qy      480  eGluGlyAlaThrGluGluGluLysLeuLeuSerArgLysMetMetLysTyTTPAlaTh
Db      1497  CGAAGGACCCACGAGGAGGAGAAATTACTAGCCCGGAAAGATGATGAATATCTGGGCTAC
Qy      500  rPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTTPProAlaTyAsnLe
Db      1557  CTTTGCTCGAAACCGGAATCTTAATGGGAAACGACCTGTCTGTGGCCAGCTTATTAATCT
Qy      520  uThrGluGlnLysLeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLysGluPr
Db      1617  GACTGAGAGTACTGACGCTGCACTTAACATGAGCTCGGACAGAGACTCAAAAGAAC
Qy      540  cArgValaAspPheThrThrSerThrLileProLeuLileuSerAlaSerAspMetLeuH1
Db      1677  GCGGGTGGAGTTTGGACCAACACCAATCCCTGATCTGTGCTGCCAGCATGCTCCA
Qy      560  sSerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePhePheCysAl
Db      1737  CAGTCCCTTTCTTCTTAACTTCTCTCTCTCTCTCCAGCTTCTTTCTTTTGTGTC
Qy      580  aPro 581
Db      1797  TCCCT 1800

```

```

RESULT 10
LOCUS   BC069548                1896 bp    mRNA    linear    PRI 30-JUN-2004
DEFINITION Homo sapiens hypothetical protein FLJ31547, mRNA (cdna clone
ACCESSION BC069548
VERSION   BC069548.1   GI:46854430
KEYWORDS  MGC.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1896)

Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Helel, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stopleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Topchiyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Boeck, S.A., McKen, P.J., McKernan, K.U., Malek, U.A., Gunaratne, P.H., Richards, S.W., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, B.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1896)

Strausberg, R. Direct Submission

Submitted (29-APR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomes Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: gcaps-r@mail.nih.gov

Tissue Procurement: Baylor Human Genome Sequencing Center

cDNA Library Preparation: Baylor Human Genome Sequencing Center

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HQSC

Web site: <http://www.hqsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulik, S.W., Louesged, H., Kowls, C.R., Sneed, A.J., Martin, R.G., Muzny, D.W., Nanavati, A.N., Gibbs, R.A.

FEATURES

source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNU at: <http://image.llnl.gov>

Series: IRBR Plate: 4 Row: F Column: 1.

Location/Qualifiers

1..1896

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:97193 IMAGE:7262439"

/tissue_type="PCR rescued clones"

/clone_lib="NIH MGC 244"

/note="Vector: pPCR-Script Amp SK(+)"

1..1896

/gene="FLJ31547"

/db_xref="LOCUSID:221223"

120..1697

/gene="FLJ31547"

/codon_start=1

/product="hypothetical protein FLJ31547"

/protein_id="AAH69548.1"

/db_xref="GI:46854431"

/db_xref="LOCUSID:221223"

/translation="MSGNWVHGGIILIAIIVLAAATPKGPSAEGPDRNTRLGWIOGKO VIVLGSPIVAVPLGVPRAAPLGLSIPRTNQDPSAMDNRBAATSYPLCQNSEWLL LDQHLKHYPKFVSEDCLYLNTIYAPHADTSGKLPLVWVFGAGRTGASIFDSD ALAAVEDLVAVVQIRLGIFFGFTTWDAFGNNAFQDVAAALSWQKNIIEFGDDPS

SVTFESAGAISSVSLILSPMAKGLPHKAIINESGVALIPYLEANDYEKSEDLQVVAH
FCGNNAASEALRLCLRTKPSKELTSLQKTSFTRVNDGAFRPEPDLISDOKAFKA
IPSIIVANNHEGFLPLPMKEAPRLISGSKSLALIONILHI PROYLIVANEYEH
KSLTRERDSLDLDGIVFVVRPALITARYNHEGATBEKLSRQMKYATPATGN
PNDUSLMPAINLTQYLDLDNMSLDQRLKRPVDFTWTSTPILISDSMDLHSPLS
SLTFLSLQPPFFFCAP

ORIGIN

Alignment Scores:

Pred. No.: 6,356-212 Length: 1896
Score: 2606.00 Matches: 503
Percent Similarity: 88.77% Conservative: 3
Best Local Similarity: 88.25% Mismatches: 6
Query Match: 84.64% Indels: 58
DB: Gaps: 2

US-10-023-515-2 (1-581) x BC069548 (1-1896)

QY 12 TrpCyvPhePheLeuIleuGlnProLeuLeuGlnIleuArgGlnTrpGlyLeuThrGly 31
DB 159 TGGGCTATCTGGCTCTTGCACGCC-----ACCAAGG 194
QY 32 ProSerAlaGluGlyProGlnArgAnthArgLeuGlyTrpIleGlnGlyLeuVal 51
DB 195 CCTTCTGCGAAGGGCCACAGAGAACACAGGCTGGATTCAGGGCAAGCAAGTC 254
QY 52 ThrValLeuGlySerProValProValaenValPheLeuGlyValProPheAlaLarPro 71
DB 255 ACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTCCTCGAGTCCCTTCTCTCTCC 314
QY 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArg 91
DB 315 CCGCTGGGATCCCTCGATTACGAACCGGACGCTGCATCCCTGGGATTAATTGCGA 374
QY 92 GluAlaThrSerTrpProAsnLeuCyLeuGlnAsnSerGluTrpLeuLeuAspGln 111
DB 375 GAAGCACTCTCACTTAATTTGCTCCAGAACTCAGATGGCTGCTTAAATCA 434
QY 112 HisMetLeuLeuValHisTrpProLeuPheGlyValSerGluAspCysLeuTrpLeuAsn 131
DB 435 CACATGCTCAAGGTCAATTACCCGAATTCGAGGTCTCAGAACTGCTCTTACCTGAAC 494
QY 132 IleTrpAlaProAlaHisAlaAspThrGlySerLeuPheProValIleTrpPhePro 151
DB 495 ATCTATGGCCCTGCCACGCCGATACAGGCTCAAGCTCCCGTCTTGCTGCTTCCCA 554
QY 152 GluGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIleTrp 171
DB 555 GAGGTGCTTCAAGACTGGCTCAGCTTCATCTTGATGGGTCCGCCCTGCTGCTAT 614
QY 172 GluAspValLeuValValValGlnTrpArgLeuGlyIlePheGlyPhePheThrThr 191
DB 615 GAGGACGTGCTGGTGTGCTGCTCAGATACCGGCTAAGAAATTTGGTTTCTTCAACCA 674
QY 192 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLeuAspGlnValAlaAlaLeuSerTrp 211
DB 675 TGGATTCAGCATGCTCCGGGAACCTGGGCTTCAAGAACAGAGCTGCTGCTGCG 734
QY 212 ValGlnLeuAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGlu 231
DB 735 GTCCAGAAAGAACATGAGTTCTTGGTGGGACCCACAGTCTGTACCACTTTTGGCGAG 794
QY 232 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLeuGlyLeuPhe 251
DB 795 TCCGGGGAGGCAATAGTGTCTAGTCTTATACGTCTCCCATGGCCAAAGTTTATTC 854
QY 252 HisLeuAlaIleMetGluSerGlyValAlaIleIleProGlyLeuGlnAlaHisAspTrp 271
DB 855 CACAAAGCATATAGAGAGTGAGGTGGCCATCATCTTACTTGAAGCCCATATATAT 914
QY 272 GluLeuSerGluAspLeuGlnValValaHisPheCysGlyAsnAsnAlaSerAspSer 291
DB 915 GAGAAAGATGAGGACTCGAGGTGTTGACATTTCTGTGGTAAACAATGCCGACAGCTT 974

QY 292 GluAlaLeuLeuArgCysLeuArgThrIlePheProSerIleGlyLeuLeuThrIleuSerGln 311
DB 975 GAGGCCCTCTGATGAGTGTGCTGAGCAAAACCTTCAAGAGACTGTGACCTTCAAGCCAG 1034
QY 312 IyethrIySerPheThrArgValValaAspGlyAlaPhePheProAsnGluProLeuAsp 331
DB 1035 AAAACAAAGTCTTTCATCTGAGTGTGATGGTGTCTTCTTCTTATATGAGCTCTAGAT 1094
QY 332 LeuLeuSerGlnLeuAlaPheValAlaIleProSerIleIleGlyValaAsnAsnIleGlu 351
DB 1095 CTAATGTCTCAAAAAGCAATTAAACCAATTCCTTCATCATGTGAGTCAATTAACACAGAG 1154
QY 352 CysGlyPheLeuLeuProMetIySerGluAlaProGluIleLeuSerGlySerAsnIySer 371
DB 1155 TGTGGCTTCTCTGCTCTTATGAGAGGCTCTTGAGATCTTCAAGTGGCTTCAACAGTCC 1214
QY 372 LeuAlaLeuHisIleuIleGlnAsnIleLeuHisIleProProGlnTrpIleuHisIleuVal 391
DB 1215 CTGGCCCTCATCTGATCAAAACATCCCTGACATCCGCGCTCAGATATTGCACTTGTG 1274
QY 392 AlaAsnGluTrpPheHisAspIleHisIleSerLeuThrGluIleArgAspSerIleuLeuAsp 411
DB 1275 GCTAATGAATACTTCCATGACAAAGCACTCCCTGACTGAATTCGAGACAGTCTTGTGAC 1334
QY 412 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTrpHisAspAsp 431
DB 1335 TTGCTTGAAGAGTGTCTTGTGTGATCTTGTGATCTGACATGACAGCTCGATATACAGA--- 1391
QY 432 AlaGlyAlaProValIyPheTrpGluPheArgHisArgProGlnCyvPheGluAspThr 451
DB 1391 ----- 1391
QY 452 LysProAlaPheValIyAlaAspHisAlaAspGluValArgPheValPheGlyGlyAla 471
DB 1391 ----- 1391
QY 472 PheLeuLeuGlyAspIleValMetPheGluGlyAlaThrGluGluGluLeuLeuSer 491
DB 1392 -----GAGGAGCAGAGAGAGAGAGAGAGATTAATCTGAGC 1424
QY 492 ArgIyMetMetLeuTrpTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 511
DB 1425 CGGAAGATGATGAATATCTGGGCTACCTTGTGTGAAACGGGAATCTTATGGAACGAC 1484
QY 512 LeuSerLeuTrpProAlaTrpAsnLeuThrGluGlnTrpLeuGlnLeuAspLeuAsnMet 531
DB 1485 CTGTCTGTGTGCGCAGCTTATATCTGACTGAGCAGTACCTTCAAGCTGACATTGAACATG 1544
QY 532 SerLeuGlyGlnArgLeuLeuGluProArgValaAspPheTrpThrIleProLeu 551
DB 1545 AGCTCTGGACAGAGACTAAAGAACCGGGGGAGATTTTGGACCAAGCAACATCCCCCTG 1604
QY 552 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 571
DB 1605 ATCTGTGTGCTCTCGAAGATGTCACAGTCTCTTCTTCTTAATCTTCTCTCTCTC 1664
QY 572 LeuGlnProPhePhePhePheCysAlaPro 581
DB 1665 CTCCAGCCTTCTTCTTCTTCTTGTGTCTCT 1694

RESULT 11
AX714052 1962 bp DNA linear PAT 15-APR-2003
LOCUS AX714052
DEFINITION Sequence 736 from Patent EP1293569.
ACCESSION AX714052
VERSION AX714052.1 GI:29888980
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homosapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1

AUTHORS

Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamehika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, T.

TITLE

Full-length cDNA
Patent: EP 1293569-A 736 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)

FEATURES

Location/Qualifiers

1..1962
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	6,64e-212	Length:	1962
Score:	2606.00	Matches:	503
Percent Similarity:	88.77%	Conservative:	3
Best Local Similarity:	88.25%	Mismatches:	6
Query Match:	84.64%	Indels:	58
DB:	6	Gaps:	2

US-10-023-515-2 (1-581) x AK714052 (1-1962)

```

OY      12  TTPCyvRheRheLeuLLeuGInProLeuLeuGInHiaRgIntrpGlyuRthrgly  31
DB      189 TGGGGATATGCGGCTCTTGACAGCCCC-----ACCAAGG  224

OY      32  PPSRRAlAGInGlyProGInaRgAntrpRgLeuGlyTTPILGInGlyuRgInVal  51
DB      225 CTTTGCTGAAGGGCCACAGAGAACACACAGCTGGATGATTCAGGGCAAGCTC  284

OY      52  ThrValLeuGlySerProValProValaRheLeuGlyValProRheAlaPro  71
DB      285 ACTGTCGCGGAAGCCCTGCTGCTGAACGTGTTCTCGAGAGTCCCTTGCTGCTCC  344

OY      72  ProLeuGlySerLeuRgRheThraRProGInProAlaSerProTTPAspRheLeu  91
DB      345 CCGCTGGGATCCCTCGATTTCAGAACCCGACGCTGCATCGCCCTGGGATACCTTGA  404

OY      92  GIuaIaThrSerTyrProAlaRheLeuCyLeuGInaRserGIntrpLeuLeuRgIn  111
DB      405 GAAGCCACCTCTACCTTAATTGCTGCCAAGATCAGATGCTGCTTGAATCAA  464

OY      112 HisMetLeuLyRValHisTyrProLyRbRheGlyValSerGIuaRProCyLeuTyrLeu  131
DB      465 CACATGCTCAAGGTGCATTACCCGAATTCGGAAGTGCAGAAAGCTGCTTACCTGAAC  524

OY      132 IleTyrAlaProAlaHisAlaAspThrGlySerLyLeuProValLeuValTTPRhePro  151
DB      525 ATCTATGGCTGCTGCCACAGATACAGGCTCCAAAGCTCCCGCTTGCTGCTCCCA  584

OY      152 GIyGlyAlaRheLyRThrgLySerAlaSerTLeRheAspGlySerAlaLeuAlaTyr  171
DB      585 GGAAGTGCCTTCAAGACGCTGCATGCTCCATCTTGAATGGGTCGCGCTGCTGCTAT  644

OY      172 GIuaRValLeuValValValGInTyrRgLeuGlyLeRheGlyRheRheThr  191
DB      645 GAGACCTGCTGCTGCTGCTGCTCAAGTACCGGCTAGGAATATTGGTCTTCAACACA  704

OY      192 TRAspRGIhNValaProGlyAAspTTPAlaRheLyRAspRGIhNValaAlaLeuSerTTP  211
DB      705 TGGGATACGACGTGCTCGGGAACTGGGCTTCAAGACCAAGGTGCTGCTGCTGCTG  764

OY      212 ValGInLyRAspNileGlyRheGlyGlyAspProSerSerValThrIleRheGlyGly  231
DB      765 GTCCAGAAAGACATCGATCTTCGAGGGGACCCAGCTCTGACCACTTTGGGAG  824

OY      232 SerAlaGlyAlaIleSerValSerSerLeuLLeuSerProMetAlaLyRGIyLeuPhe  251
DB      825 TCCGCGGAGCCATAGTGTTCATAGCTTATACGTCTCCCATGCGCAAGGCTTATTC  884

```

```

OY      252 HisLyRValaIleMetGlySerGIyValAlaIleIleProTyrLeuGIuaLHisAspTyr  271
DB      885 CACAAGCCATCATAGAGAGTGGGGTGGCCATCATCCCTTACCTGGAGGCCCATGATTAT  944

OY      272 GIuLySerGIuaRLeuGInValValaHisRheRGIyAAspNAlaSerAspSer  291
DB      945 GAAAGATGAGAACCTTCAGAGTGGTGCACATTTCTGTGGTAACTATGCTCGACACTCT  1004

OY      292 GIuaLLeuLeuAArgCyLeuAArgThrLyRAspSerLyRGIyLeuLeuThrLeuSerGI  311
DB      1005 GAGGCCCTGCTGAGGTGCTGAGGACAAACCTCCAAAGAGCTGCTGACCTTACGCCAG  1064

OY      312 LyRThrLyRSerRheThraRValValaAspGIyAlaRheRheProaRGIuProLeuAsp  331
DB      1065 AAAACAAAGTCTTCACTCGAGGTGATGCTCTTTCTTCTTAATGAGCCTCTAGAT  1124

OY      332 LeuLeuSerGInLyRAlaRheLyRAlaGleProSerTLeIleGlyValaAspNHisGly  351
DB      1125 CTATTGCTCGAAGAGCATTTAAAGCAATTCCTTCCATTATCGAGTCAATACACAGAG  1184

OY      352 CyGlyRheLeuLeuProMetLyRGIyAlaProGIuLLeuSerGIySerAspLySer  371
DB      1185 TGTGGCTTCCCTGCTGCTTATGAAGAGGCTCTGAGATCCCTGAGTGGCTCCAAACAGTCC  1244

OY      372 LeuAlaLeuNHisLeuLileGInaRNIleuNHisIleRProRGIuTyrLeuNHisLeuVal  391
DB      1245 CTTCCCTCCATCTGATTAACAACATCTGCATCCCGCTCAAGTATTTGACACTTGTG  1304

OY      392 AlaAspGIuTyrRheNHisAspLyRHisSerLeuThrgInLileRAspSerLeuLeuAsp  411
DB      1305 GCTAATGATATCTTCATGACAAAGCACTCCCTGATGAAATCCGAGACAGTCTTCTGGAC  1364

OY      412 LeuLeuGlyAspValaRheRheValaProAlaLeuLileThraLaaRgTyrHisAspAsp  431
DB      1365 TTGCTTGAAGATGCTGCTTGTGATCCCTGACATGATCAAGCTGATATCAGAGA---  1421

OY      432 AlaGIyAlaProValTyrRheTyrGIuRheRgHisAspProGInCyRbRheGIuaRThr  451
DB      1421 -----  1421

OY      452 LyRProAlaRheValLyRValaAspNHisAlaAspGIuValaRbRheValaRheGlyAla  471
DB      1421 -----  1421

OY      472 RheLeuLyRGIyAspRileValMetRheGlyGlyAlaRThrgInGlyGlyLyLeuLeuSer  491
DB      1422 -----GAAGAGCCACGAGAGGAGAGAAATTACTGAGC  1454

OY      492 ArgLyMetMetLyRtTTPAlaThrRheAlaArgThrgLyAspProAspGIyAspAsp  511
DB      1455 CGGAAGATGATGAATACTGGGCTACCTTGTCTCGAACCAGGAATCTTAATGGAAAGCA  514

OY      512 LeuSerLeuTTPProAlaTyrAspLeuThrgInGlyTyrLeuGInLeuAspLeuAspMet  531
DB      1515 CTGTCTCTGTGGCACCTTAATCATCTGAGCAGCTACCTCAGCTGAGTGAACATG  1574

OY      532 SerLeuGlyGInaRgLeuLyRGIuProAArgValaAspRheTTPThrSerThrIleProLeu  551
DB      1575 AGCTTCGACAGAGACTCAAAAGACCGCGAGTGAATTTTGAACAGACCAATCCCCCTG  1634

OY      552 IleLeuSerAlaSerAspMetLeuNHisSerProLeuSerSerLeuThrRheLeuSerLeu  571
DB      1635 ATCTGTCTGCTCGACATGCTCCACAGTCTCTTCTTCTTAACCTTCTCTCTCTC  1694

OY      572 LeuGInProRheRheRheRheCyRAlaPro  581
DB      1695 CTCACGCTTCTTTCTTTGTGCTCTCT  1724

```

RESULT 12

AK056109 1962 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ31547 f16, clone NT2R12001010, weakly similar
DEFINITION to PATTY ACYL-CoA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).

ACCESSION AK056109
VERSION AK056109.1 GI:16551422
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS
1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakematsu, A., Hayashi, K., Satoh, H., Nagai, K., Kimura, K., Nakita, H., Sekine, M., Obaishi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahata, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, T., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirooka, S., Chiba, Y., Iehida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Hara, H., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsumura, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kunikida, A., Itakura, S., Fukuzumi, T., Fujimori, Y., Komiyama, M., Tashiro, H., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Ozaki, K., Hirose, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Sato, H., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Maehuo, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isozaki, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
2 Nishimura, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakematsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahata, K., Masuno, Y., Nagai, K. and Isozaki, T. NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 1962)
Isozaki, T., Otsuki, T. and Sugiyama, T. Direct Submission
Submitted (24-OCT-2001) Takao Isozaki, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5- & 3-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
Source
1. 1962
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2R12001010"

CDS
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2R12"
/note="Cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction. -majorly NT2 neuron 150. 1727
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71094.1"
/db_xref="GI:16551423"
translation="MSGMWVHPGOILLWAIWLAAPTKPSAEGPQNRTRIGWISQKQ
/TVVIGSPVAVVFLGVPAPAPLGLSRLFTNQPSAPMDNLEATSPYKICONSFML
LDOHMLKYHPKFGVSEDCYLYNTPAPADPGSKLPVLPWPGGAPKFGSAKIDGS
ALAAVEDLVVVVOYRGLDIPGFPTWDOHAPGNAPKOVVALASVOKNIEFGSDPS
SVTFGEBSAGAISVSSSLISPMAGKLPKALMESSVALITPLBAIDYKSESDQYVAH
FCGNASDSEALRLRLTRKPESELTSOKTSFTRVVDGAFPPNEPLDLSQKAFKA
IPSIIVGNVHSCGFLPMKEAPPELISGNSKSLALHIONLHPIPOYHLVANEYFD
KSLTEIISDLSLDLDGVFVVPALITARYHREGATEEKLISRMQKYMVFARTGN
PNGNDLSLMPVYNLTROYLQDLNMSLGRLKEPRVDVMTSTIPLISAPDLHSPLS
SLTFLSLQPPFPFCAP"

ORIGIN
Alignment Scores:
Pred. No.: 6,646-212 Length: 1962
Score: 2606.00 Matches: 503
Percent Similarity: 88.77% Conservative: 3
Best Local Similarity: 88.25% Mismatches: 6
Query Match: 94.64% Indels: 58
Gaps: 2

US-10-023-515-2 (1-581) x AK056109 (1-1962)

QY 12 TTPCySPhePheLeuIleLeuGlnProLeuLeuGlyHISArgGlnTTPGlyVlySerThrGly 31
DB 189 TGGGCTATCTGGTGGCTCTGACAGCCCC-----ACCAAGGG 224

QY 32 ProSerAlaGluGlyProGlnArgGlnThrArgLeuGlyTTPILGlnGlyLeuGlnVal 51
DB 225 CCTTGTGCTGAAGGGCCACAGAGAACCCAGCTGGAGATTGAGGCGAAGCAAGATC 284

QY 52 ThrValLeuGlySerProValProValaAnValPheLeuGlyValProPheAlaIaPro 71
DB 285 ACTGTGCTGGAGAGCCGCTGCTGTGAACGTTTCTCGAGAGTCCCTTGTCTCTCTCC 344

QY 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeuArg 91
DB 345 CCGCTGGGATCCCTCGATTTCGAACCCGACGCTGATCGCCCTGGGATTAATTGGCA 404

QY 92 GluAlaThrSerTyrProAsnLeuCyLeuGlnAsnSerGlnTTPLeuLeuAsnArg 111
DB 405 GAAGCCACCTCTTACCTTAATTTGTGCTCCGAACCTCAAGATGGCTGCTTAATCA 464

QY 112 HisMetLeuValValHisTyrProLysPheGlyValSerGluAspCyLeuTyrLeuAsn 131
DB 465 CACATGCTCAAGTGCATTACCCGAATTCGAGAGTGTCAAGAAATGCTCTTACCTGAAC 524

QY 132 IleTyrAlaProAlaHisAlaAspThrGlySerTyrLeuProValLeuValTTPPhePro 151
DB 525 ATCTATGGCGCTGCCACGCGATTCAGAGCTTCAAGCTCCCGTGTGTGGTGTCCCA 584

QY 152 GlyGlyValaPheLysThrGlySerAlaSerTyrPheAspGlySerAlaLeuAlaIaTyr 171
DB 585 GAGAGTGCCTTCMAACGCGCTCAAGCTTCATCTTGAATGGTCCGCGCTGCTCTCAT 644

QY 172 GluAspValLeuValValValGlnTyrArgLeuGlyTyrIlePheGlyPhePheThrThr 191
DB 645 GAGGACGTGCTGT 704

QY 192 TTPAspGlnHisAlaIaProGlyAsnTTPAlaPheLysAspGlnValAlaAlaLeuSerTTP 211
DB 705 TGGATTCAGCATGCTCCGGGGAACCTTCAGGACCAAGTGTGTGTGTGTGTGTGTGT 764

QY 212 ValGlnLysAsnIleGluPheGlyGlyAspProSerSerValThrIlePheGlyGln 231
DB 765 GTCAGAGAAATCGAGTCTTCGGTGGGAGCCCAAGCTCTGACCATCTTTGGCGAG 824
QY 232 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIleGlyLeuPhe 251
DB 825 TCCGCGGAGACCAATAGTGTCTTACCTTATCTGCTCCCAATGGCCAAAGGCTATTTC 884
QY 252 HisValAlaIleMetGluSerGlyValAlaIleIleProGlyLeuGlnAlaHisAspTyr 271
DB 885 CACAAAGCCATCATGAGAGAGTGGGGTGGCATATCCCTTACCTGAGGGCCATGTATAT 944
QY 272 GluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 291
DB 945 GAGAAAGTAGAGGACCTGCGAGGTGTGCACTTTCTGTGATACATCGTCAGCTCT 1004
QY 292 GluAlaLeuLeuArgCysLeuArgThrIleProSerLysGluLeuLeuThrLeuSerGln 311
DB 1005 GAGGCCCTCTGAGGTGGCTGAGGACAAACCTCCAGAGAGCTGTGACCTCAGCCAG 1064
QY 312 LysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 331
DB 1065 AAAACAAAGCTCTTTCCTCGAGAGTGTATGTGTCTTTCTTCTTATAGCCCTCAGAT 1124
QY 332 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHisGln 351
DB 1125 CTATTGTCTCAAGAAAGCATTTAAAGCATTCCTTCATATCGAGTCAATACCAAGAG 1184
QY 352 CysGlyPheLeuLeuProMetLysGluAlaProGlnIleLeuSerGlySerAsnLysSer 371
DB 1185 TGTGCTTCTCTGCTGCTATGAGAGAGCTCTGAGATCTTCAAGTGGCTCCAAAGTCC 1244
QY 372 LeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIleLeuHisLeuVal 391
DB 1245 CTTGGCCCTCACTGATACAAACATCCGACATCCGCGCTCAATATTGACCTTGTG 1304
QY 392 AlaAsnGluTyrPheHisAspLysHisSerLeuThrGlnIleArgAspSerLeuLeuAsp 411
DB 1305 GGTAAATGAAATCTTCATGACAAAGCATCCCTGACTGAATCCGAGACAGTCTTGTGAC 1364
QY 412 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAsp 431
DB 1365 TTGCTTGAGAGATGCTTCTTGTGTGCTCCCTGACATGATCAAGCTCGATATCAAGA--- 1421
QY 432 AlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAspThr 451
DB 1421 ----- 1421
QY 452 LysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyAla 471
DB 1421 ----- 1421
QY 472 PheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGlnLysLeuLeuSer 491
DB 1422 -----GAAGAGCCACGAGAGAGAGAAATTTACTGAGC 1454
QY 492 ArgLysMetMetLysTyrTTPAlaIlePheAlaArgThrGlyAsnProAsnGlyAsnAsp 511
DB 1455 CGAAGAGATGATGAATACTGGGCTAACCTTGTCTGAAACCGGGAATCTTAATGGGAAGAC 1514
QY 512 LeuSerLeuTTPProAlaTyrAsnLeuThrGluGlnIleLeuGlnLysAspLeuAsnMet 531
DB 1515 CTGCTCTGTGGCCAGCTTATATCTGACTGACGATACCTCCAGCTGGACTTGAACATG 1574
QY 532 SerLeuGlyGlnArgLeuLysGluProArgValAspPheTTPThrSerThrIleProLeu 551
DB 1575 AGCCTCGGACAGAGACTCAAGAACCGGGGTGATTTTGGACCAACACATCCCTG 1634
QY 552 IleLeuSerLysSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 571
DB 1635 ATCTGTGCTGCTCGACATGCTCCACAGTCTCTTCTTCTTAACCTTCTCTCTC 1694
QY 572 LeuGlnProPhePhePhePheCysAlaPro 581

DB 1695 CTCAGACCTTCTCTTCTTCTTGTGCTCT 1724
RESULT 13
AX958392 1857 bp DNA linear PAT 14-JAN-2004
LOCUS AX958392
DEFINITION Sequence 20 from Patent WO0226588.
ACCESSION AX958392.1 GI:40879350
VERSION AX958392.1
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Human drug metabolizing enzymes
Patent: WO 0226988-A 20 04 -Apr-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
location/Qualifiers
1..1857
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,91e-210 Length: 1857
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 84.07% Indels: 23
DB: 6 Gaps: 1
US-10-023-515-2 (1-581) x AX958392 (1-1857)
QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTTPriIleGlnGlyLysGln 50
DB 274 GGGCTTCTCTGTAAGGCGCACAGAGAACACAGCTGGAGTGGATTCCAGGGCAGACAA 333
QY 51 ValThrValLeuGlyLysProValProValAsnValPheLeuGlyValProPheAlaIle 70
DB 334 GTCACTGTCTGGAGAGCCCTGTGCTGAGAGTGTCTCGAGTCCCTTGTGCTGCT 393
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeu 90
DB 394 CCCCCGTGGATCCCTTGCATTTACGAACCCGACCTGCATCGCCCTGGGATTACTTG 453
QY 91 ArgGluAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGluTTPLeuLeuAsp 110
DB 454 CGAAGAGCCACCTCTTAACCTTAATTTGTGCTCCAGACTCAAGAGTGGCTCTTAAGAT 513
QY 111 GluHisMetLeuLysValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeu 130
DB 514 CAACATATGCTCAAGGTGATTAACCCGAATTCGAGGTCTCAGAAACTGCTCTACCTG 573
QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTTPPhe 150
DB 574 AATCATTAATCGCGCTCCCAAGCCGATACAGAGCTCCAAAGCTCCCGCTCTTGATGTG 633
QY 151 ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIle 170
DB 634 CCAGAGGTGGCTTCAAGACTGAGCTCAGCTCATCTTGATGGGATCCGCTGGGTGCG 693
QY 171 TyrGluAspValLeuValValValGlnTyrArgLeuGlyTlePheGlyPhePheThr 190
DB 694 TATGAGAGAGT 753
QY 191 ThrTTPAspGlnHisAlaProGlyAsnTTPAlaPheLysAspGlnValAlaAlaLeuSer 210
DB 754 ACATGGAGATCAACAGAGTCTCGGGGAACTGGGCTTCAAGAGACCAAGTGGCTCTGTG 813
QY 211 TTPValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 230

Db 814 TGGGTCAGAGAAATCATGATTTCTTCGGTGGGAGCCCGAGCTGTGACCATCTTTGGC 873
Qy 231 GUSERAGAGYALAIIESEVALSERSELEUUIELEUSERPROMETALALYSGLYLEU 250
Db 874 GAGTCCGGGGGAGCAATAGTGTCTTACTTTATACGTCTCCATGGCCAAAGGCTTA 933
Qy 251 PHEHISLYSALAIEMERGUSERGIVAIAlAIIELEPROTYRLEUGIUAHIAAP 270
Db 934 TTCCACAAAGCCATCATGAGAGAGTGGGGGCGCCATCATCTTACTTGAGGCGCCATGAT 993
Qy 271 TYRGIVUSSESGIUSAPLEUGIIVALAIAlAHIPHECYEGIYASNAANAIASEAEP 290
Db 994 TATGAGAGAGTGAAGACCTGCAGAGGTGTTCACATTTCTGTGTAACAATGCGTCAGAC 1053
Qy 291 SERGIUALALEUUAARGCYALEUAAGTHIRLYRPROSERILEIISGLIYALEUETHIRLEUSER 310
Db 1054 TCTGAGGCCCTGCTGAGGAGCTGAGGACAAACCTCCAAAGAGCTGCTGACCCCTGAC 1113
Qy 311 GINLYETHIRLYSERPHETHRARGVALAIAPGIVAIAPHEPROANGIUPROLEU 330
Db 1114 CAGAAACAAAGCTTTTCACTGAGGTGTGATGGTCTTTCTTCTTATAGCCTTA 1173
Qy 331 ASPLEUSEUSERGILNLYSALAPHEUYALAIIEPROSERILEIISGLIYALASNAHIS 350
Db 1174 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGTCAATACCAAC 1233
Qy 351 GIUCYSGIYPHLEULEUPROMETLYSGIUALAPROGLIUILEUSERGILSERHNLVS 370
Db 1234 GAGTGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
Qy 371 SERIEUALALEUHIIELEUILEGIMENIIELEUHSILEPROPROGLINTYRLEUHSILEU 390
Db 1258 -----CACATCCCGCTCAGATTTGACCTT 1284
Qy 391 VALAIASANGIUTYRPHENISAPLYSHISERLEUTHIRGILIEARGHAPSERLEU 410
Db 1285 GTGGGTATGATATCTTCATGACAGACATCCCTGACATGAAATCGAGACAGTCTTCG 1344
Qy 411 ASPLEULEUGIYAPHEVALAPHEPHEVALAIIPROALALEUIETHRALAAGTGYHISARG 430
Db 1345 GACCTGTGGAGATGTGTTCTTTGTGGTCCCTGACATGATCAGCTCATATACAGAA 1404
Qy 431 ASPRAGIYALAPROVALTYRPHETRYRGIPHEAGTHISARGPROGLINCYAPHEGUAAP 450
Db 1405 GATGCTGGGACCGCTGCTACTTCTATGAGTTTGGCACCGGCTCAGTGTGTAAGAC 1464
Qy 451 THIRLYSPROALAPHEVALIYSAALASPHISAIAPSGIUALARGPHEVALAPHEGLIY 470
Db 1465 ACGAAGCCAGCTTTGTCTCAAAACCCAGCCGCTGATGAAAGTCCGCTTGTGCTGGTGT 1524
Qy 471 ALAPHELEUYEGIYAPRIIEVALMECPHEGUGIYALATHRGINGIUGIULYSLEU 490
Db 1525 GCCTTCTGGAAGGGGACATTTGTATGTTGGAAGGACCGAGAGGAGGAGAAAGTTACTG 1584
Qy 491 SERATGIVAMEMETLYETTRPALATHRPHEALIAARGTHRGIYASMPROANGIYAN 510
Db 1585 AGCCGGAAGATGATAAATACTGGGCTACCTTGTCCGAACCGGAATCTTAATGGGAAC 1644
Qy 511 ASPLEUSERLEUTRPROALATYRANLEUTHIRGILNLYRLEUGIULEUASPLEUAN 530
Db 1645 GACCTGTCTCTGTGGCCAGCTTATATCTGACTGAGCAATCACTCAAGTGAAGCTTGAAC 1704
Qy 531 METSERLEUGIYGLARGLEUYSGIUPROARGVALASPPHETRTPHSETHIRILEPRO 550
Db 1705 ATGAGACCTCGACAGAGACTCAAGAAACCGCGAGAGAGATGTGTGGGTGACGGGGTATCCT 1764

RESULT 14
CQ733044 1716 bp DNA linear PAT 03-FBB-2004
LOCUS CQ733044
DEFINITION Sequence 18978 from Patent WO02068579.
ACCESSION CQ733044
VERSION CQ733044.1 GI:42314969

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.
TITLE Kite, such as nucleic acid arrays, comprising a majority of humenexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 18978 06-SEP-2002;
FEATURES
source location (NY) (US)
1. 1716
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:

Pred. No.:	2,2e-203	Length:	1716
Score:	2505.00	Matches:	478
Percent Similarity:	93.98%	Conservative:	6
Best Local Similarity:	92.82%	Mismatches:	15
Query Match:	81.36%	Indels:	16
DB:	6	Gaps:	1

US-10-023-515-2 (1-581) x CQ733044 (1-1716)

Qy 31 GIYPROSERIAGIYALIGIYPROGINARGANTRARGLEUGIYTRPIEGINLYYSGIN 50
Db 28 GGGCTTCTGCTGGAAGGGGCAAGAGAAACAGAGGCTGGATGATCAGGGCAAGCA 87
Qy 51 VALTHRVALLLEUGIYSERPROVALPROVALAENVALPHELEUGIYVALPROPHENIAAIA 70
Db 88 GTCACGTGTGCTGGGAAGCCCTGTGCTGGAACGTGTCTTGAGAGTCCCTTGTGCT 147
Qy 71 PROPLEUUGIYSELEUAPHETHRANPPOGINPROALASERPROTRPAASPLEU 90
Db 148 CCCCCTGGAGATCCCTGCAATTTAGAACCCGACGCTGCAATCGCCCTGGATTAATCTG 207
Qy 91 ARGIVUALATHSERTYRPROANLEUCYSEUGIANSERGIUTRPLEUENLEUAP 110
Db 208 CGAGAGGCACTCTTACCTTAATTTTGCTCCAGAACTGAGATGCTGCTTAAAT 267
Qy 111 GINIEMETLEUYVALHISTRYPROLYSPHEGIVALSERGIUNAPCYSEUTYRLEU 130
Db 268 CAACACATGCTCAAGGTCATTACCCGAAATTCGAGAGTCAAGAGACTGCTTAACCTG 327
Qy 131 AENIETRYRALAPROALAHISALASPHIRHISERLYSEUVALLEUVALTRPHE 150
Db 328 AACATCTATGGGCTGCCACCGCCGATACAGCTCAAGCTCCGCTTGTGGTGTTC 387
Qy 151 PROGIYIYALAPHELYETHRGIYSEALASERILEPHEAPGLYSERALALEUALAIA 170
Db 388 CCAGAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAAGGATCCGCTTGGCTGCC 447
Qy 171 TYRGIVAPVALLEUVALIYVALIYALGINTYRARGLEUGIYIAPHEGILYRPHETHR 190
Db 448 TATGAGAGCGTGTGCTGTTGGTGTGCTCAAGTACCGGCTTAAGAAATTTGGTTCTTCA 507
Qy 191 THRTPAARGINHAIALAPROGIYANTRPALAPHELYASPGIIVALAIAlALEUSER 210
Db 508 ACATGGAGATCAGACAGTCTCCGGGGAACCTGGCTTCAAGACAGAGTGGCTGTCTGCC 567
Qy 211 TRPVAGINLYSASNIIEGLUPHEPHEGLIYASPPROSERVALTHRIIEPHEGLIY 230
Db 568 TGGGTCCAGAGAAACATGAGATCTTCCGTGGGAGCCCAAGCTCTGATACATCTTTGGC 627
Qy 231 GUSERAGIYALAIIESEVALSERSELEUUIELEUSERPROMETALALYSGLYLEU 250
Db 628 GAGTCCGGGGGAGCATAAAGTGTCTTACTGCTTATATGCTCCATGACCAAGGCTTA 687

QY 251 PheHisIySaIaIleWecGluSerGlyValaIaIeIleProTyrLeuGluIaHisAap 270
 Db 668 TTCCACAAAGCCATCATGAGAGTGGGGGCGCATCATCTTACCTGAGGCCATCATGAT 747
 QY 271 TyrGluYSerSerGluAapLeuGluValaIaHisApeCySeGlyAAsnaanaIaSerAap 290
 Db 748 TATGAGAAAGAGTGAAGACCTGACAGGTGGTTGACATTTCTGTGTAAACAATCGTCAGAC 807
 QY 291 SerGluAaIaLeuLeuAargCyLeuAargThrThyLysProSerIyGluLeuLeuThrLeuSer 310
 Db 808 TCTGAGGCCCTGTCTAGAGTGGCTGAGAGCAAAACCTCCAAAGACCTGCTGACCTCAGC 867
 QY 311 GluYThrIySerSerPheThzArgValaIaAapGlyAaIaPhePheProAmsGluProLeu 330
 Db 868 CAGAAACAAAGTCTTTCACCTGAGTGGTGAATGGTCTTCTTCTTAATAGCTCTTA 927
 QY 331 AapLeuLeuSerGluIyAaIaPheIySaIaIleProSerIleIleGlyValaIaAAsnaHis 350
 Db 928 GATCATTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 987
 QY 351 GluCySeGlyPheLeuLeuProMetIySeGlyAaIaProGluIleLeuSerGlySerAmsIy 370
 Db 988 GAGTGTGCTTCCTGCTGCTCTTAAGAGGCTCTGAGATCTCAGTGGCTCCAAACAG 1047
 QY 371 SerIeuaIaLeuAaIeLeuIleGluAAsnaIleLeuHisIleProProGluIyThrLeuHisIleu 390
 Db 1048 TCCCTTGGCCCTCCATCTGATCAAAACATCCCTGCAATCCCGCCCTCAGATTTTGACCTT 1107
 QY 391 ValaIaAaengIyTyrPheHisAapIyHisIleSerLeuThrGluIleAargAapSerLeuLeu 410
 Db 1108 GTGGCTAATGAAATACCTCCATGACAAAGACCTCCGACATCGAATCCGAGACAGTCTTCTG 1167
 QY 411 AapLeuLeuGluIyAaPvaIaPhePheValaIaProAaIeIleThrAaIaArgTyrHisAarg 430
 Db 1168 GACTTGCTTGAAGATGATGTCTTCTTGTGGTCCCTGACATGATCACAGCTCGAATATACAGA 1227
 QY 431 AapAaIaGlyAaIaProValaIyTyrPheTyrGluPheAargHisAargProGluIyCyPheGluAap 450
 Db 1228 GATGCTGGTGACCGCTGCTACTTCTATGAGTTTGGGACCGGCCCTCAATGCTTTGAAGAC 1287
 QY 451 ThrIySProAaIaPheValaIyAaHisIleAaAapGluValaAargPheValaIaPheGlyIy 470
 Db 1288 ACGAAGCGGCTTTGTCAAAGCCGACACCGCTGATGAAGTCCGCTTTGTTCGTGCTGT 1347
 QY 471 AlaPheLeuIySgIyAaPvaIaIeValaIaMerPheGluGluIaIaThnGluIyLeuLeuLeu 490
 Db 1348 GCCTTCCGAAAGGGGACATTTATGTTCCAAAGAGCCACGAGAGAGAAAGTTACTG 1407
 QY 491 SerAargIyMeMetIyTyrTyrPalaIaThrPheAaIaArgThrGlyAAsnProAAsnGlyAAsn 510
 Db 1408 AGCGGAAAGATGAGAAATATCTGGGCTACTTGTGTCGAACCGGACG----- 1455
 QY 511 AapLeuSerLeuTyrProAaIaTyrAAsnLeuThrGluIyTyrLeuGluIleuAapLeuAAsn 530
 Db 1456 -----CACTTTCATATCAGAGAGAGTGTCT 1479
 QY 531 MetSerLeuGluIyAargLeuLeuGluProAargValaAapPheTyr 545
 Db 1480 GTGATGATTAACCAAAATGCAAGTCCCAAGCCCAACCCCACTTGG 1524
 RESULT 15
 LOCUS AB186392 1728 bp mRNA linear MAM 06-AUG-2004
 DEFINITION Canis familiaris cauxin mRNA for carboxylesterase-like urinary
 ACCESSION AB186392
 VERSION AB186392.1 GI:51014274
 KEYWORDS
 SOURCE
 ORGANISM
 Canis familiaris (dog)
 Canis familiaris
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 REFERENCE
 1

AUTHORS Miyazaki, M., Yamashita, T., Taira, H. and Suzuki, A.
 TITLE Published family protein
 JOURNAL Published only in Database (2004)
 REFERENCE 2 (bases 1 to 1728)
 AUTHORS Miyazaki, M., Yamashita, T., Taira, H. and Suzuki, A.
 TITLE Direct Submission
 JOURNAL Submitted (04-AUG-2004) Masao Miyazaki, RIKEN, Frontier System
 Research, Supra-Biomedical System Research Group, Hirosewa 2-1,
 Wako-shi, Saitama 351-0198, Japan
 (E-mail: mmiyazaki@postman.riken.jp, Tel: 81-48-467-9619,
 Fax: 81-48-467-9620)
 FEATURES
 source
 1..1728
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /chromosome="2"
 1..1728
 /gene="cauxin"
 /gene="cauxin"
 1..1728
 /gene="cauxin"
 /codon_start=1
 /product="carboxylesterase-like urinary excreted protein"
 /protein_id="BAD35015.1"
 /db_xref="GI:51014275"
 /translation="MSGEMGHIGQTLTAVWVLAAREGPAADAPVPSRTLGWRGKO
 ATVUGSTMPVNVPLGITPRAPRPGRLPRKPKALLMNDSDRATSPKLCIQNSWIL
 SDQHLKVAHPNLSVEDCLYNTIYPAHANKTSKLPVWVPEGAEETSASIFDS
 ALAAVEDVLVTQYRLGIFGFKTGDQAAGWAFLDQALATWQENTLFRGDDH
 SVTIFFESAGAILSVSGIYLSPPMASGLFHKAIMSGVALIFPLRAPDERNEDIQVLR
 ICGNVSDSVALLQCLRAKSEELIDINKTKSFTRVVDGFPEPDEPLDLTEKTFNS
 IPSVIGVNNHECGFLPMPKEPRLIGGSNSLSLHLIHRVLIHPNOVLYVADQVFN
 KHSVEIRDSFLDLGDFRVVGVYARVHRVAGAVVYEEQHPPECCKOTRRPVRV
 KAHSDIRFVFGAFLKGNITVMEGATSEKLSRQWRYMAFPATGVPNGSLPL
 WPAVSQSEQYLKDLNLSVQKLEQVEPWSPTLPIMSMSTAPRGPVPLSLSVL
 LPLFFSSAP"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,136-181 Length: 1728
 Score: 2250.00 Matches: 417
 Percent Simlarity: 86.75% Conservative: 61
 Best Local Simlarity: 75.68% Mismatches: 73
 Query Match: 73.08% Indels: 0
 DB: 4 Gaps: 0
 US-10-023-515-2 (1-581) x AB186392 (1-1728)
 QY 31 GlyProSerAaIaGluGlyProGluAargAAsnThrAargLeuGlyTyrPileGluIyLeuGln 50
 Db 73 GGTCCAGCTGCTGATGACACCGGTGAGAGACACCAAGCTGGGATGGTCCGGGCGAAGCA 132
 QY 51 ValThrValleuGlySerProValaIaAsnValaIaPheLeuGlyValaIaProPheAaIa 70
 Db 133 GCCACTGTGCTGGGAGACACCATGCTGGAAGTGTCTCCGGAGATCCCTTGTGTCGC 192
 QY 71 ProProLeuGluIySerLeuAargPheThrAAsnProGluIaProAaIaSerProTyrAaPheLeu 90
 Db 193 CCCCCTTAGAGACCCCTGCGCATTTAAGAGACCAAGCATTCATTCCTGTAATGATCA 252
 QY 91 ArgGluAaIaThrSerTyrProAAsnLeuCyLeuGluIaAAsnSerGluTyrPleLeuLeuAAsp 110
 Db 253 CGAGATGCCACGTCCTTAATCTTAATGTGCTCCAGAACCTCAGTGGCTGCTCAGAT 312
 QY 111 GlnHisMetLeuIyValaHisIyTyrProLysPheGlyValaSerGluAapCyLeuTyrLeu 130
 Db 313 CAACACTTCTCAAGGTCATTAACCCCAATTTAGAAAGTCCGAAAGCTGCTGTAACCTT 372
 QY 131 AsnIleTyrAaIaProAaIaHisAaAapThrGlySerIySerIyLeuProValleuValaIaTyrPhe 150
 Db 373 AACATTAACCAACAGCCACGCGAAACACCGGCTCCAAAGTCCCTGTAATGAGTGTGCTTC 432
 QY 151 ProGluYolValaIaPheIyThrGlySerAaIaSerIlePheAapGlySerAaIaLeuAaIa 170

```

433 CCTGAGGAGGCTTTGAGACTGCTCAGCTCATCTTGATGGGCGCCCTGGCTGCC 492
QY 171 TTTGTAAPVALLEUVALVALVALGINTYRAAGLEUGLYLLEPHEGLYPHEPETHR 190
DB 493 TATGAGGATGTGCTTAATGTGCAACCCGAGTACCGGCTAAGAAATTTGGTCTTCTCAAG 552
QY 191 ThrTTPAPGSLNHIALAPROGLYASNTPLAPHELYASPGLVALAALAEUSER 210
DB 553 ACGAGGACCAAGCAAGCTCCGAGGAACTGGGCTTCTTGACCAAGTTGGCTCTTAAAC 612
QY 211 TTPVALGINTYASNLLEGLUPHEPHEGLYGLYASPProSerSerValThrILEPHEGLY 230
DB 613 TGGGTCCAGAGAAATATGAGTTCTTCCGGGGGAGCCCACTGTGTGACCATCTTGGC 672
QY 231 GLUSERLALGLYALILESERVALSERSEULEULEUSERProMECALALYGLYLEU 250
DB 673 GAGTCAGAGAGCAATAGTGTTCGGGCTGTGACTGTCCCATGGCCAGTGGCTTAA 732
QY 251 PHEHSLYVALALEMERGLUSERGLYVALAALILEPROTYRLEUGLUALAHISAPR 270
DB 733 TTCACAAAGCCATAGAGAGAGTGGGAGGCAATCATCTTCTTCCGGGCCCCGAT 792
QY 271 TYRGLIUSERSERGLUASPLeuGLINVALVALAHISAPHECYEGLYASNAVALASERAPR 290
DB 793 GATGAGAGAAATGAAATTGTCAGGTGATGCAAGTATGTGTGCAATGATACAGAC 852
QY 291 SERGLUALALEULEUARGCYEULARGTHLYSPROSERLYSGILEUSERGLYLEUThrLEUSER 310
DB 853 TCCGTAGCTGTGCTCAGTGGCTCGAGGCGCAAAATCTCCGAGAGAGTGTGTGACATCAAC 912
QY 311 GLINYSThrLYSERPHEThrARGVALVALASPGLYALAPHEProANGLUPROLEU 330
DB 913 AAGAAACCAAGCTCTTCACTGAGGTGTGAAGTTTCTTCTTCATGAGAGCTCTA 972
QY 331 ASPLeuLeuSERGLINLYSALAPHELYVALALEPROSERILEGLYVALASNAHIS 350
DB 973 GACCTATTGACGAGAGAAACATTTAATTCAATCTCTGTGATCGAGTCAATAACGAC 1032
QY 351 GLUCYSGLYPHELEUUPROMEGLYGLUALAPROGLNILEUSERGLYSEANLYS 370
DB 1033 GAGTGTGGCTTCTCTGCTCCCAAGAAAGTTCGCGAATCTTGAAGCTCCCAACAG 1092
QY 371 SERLEUALALEUHIIELEUHIIELEUHIISILEPROPROGLNTHYRLEUHIISLEU 390
DB 1093 TCTTCGCCCTGCACTGATCCACGAGCTGTCACATCCCAACCAAGTACTTAATCTT 1152
QY 391 VALALAANGLUYRPHENISAPLYSHISSErLEUTHRGLULEARGSPERLEU 410
DB 1153 GTGGCTGATCATACTTCTACAAACAGCACTCCCACTTGAGATTGAGATTTCGTG 1212
QY 411 ASPLeuLeuGLYASPVVALPHEPHEVALVALPROALALeuILEThRALARGTYRHISARG 430
DB 1213 GACTTGCTTGAGAGATGTGTTCTTGTGTGCTCTGGGGTGTGTACAGCTCGATATCACAGA 1272
QY 431 ASPALAGLYALAPROVALTYRPHETRYGLUPHEARGHISARGPROGLINCYSPHEGLUAPR 450
DB 1273 GATGCTGGTGCACCTGTCTACTTCAACGATTCAACACCACTGAGTCTTAAGAGAC 1332
QY 451 ThrLYSPROALAPHEVALYVALASPHISALASPGLUVALARGPHEVALPHEGLYGLY 470
DB 1333 ACGAGGCCAGCTTTGTCTCAAGCCGATCACTGTGATGAATACTCGCTTGTCTTGGAGGT 1392
QY 471 ALAPHELEUYEGLYASPILEVALMECPHEGLUGLYALATHRGLUGLUGLYLEULEU 490
DB 1393 GCTTCTCGAAGGCAACATGTCTCATGTTCGAAGGAGCCACAGAGAGAGAAATGGCTG 1452
QY 491 SERARGLYSPMETMETLYERTTTPALATHRPHENIALARGTHRGLYASNPROANGLYAEN 510
DB 1453 ACGAGAAAGATAGTAGGTAAGGCACTTGTCTGGAGCTGGAGCCCTAATGGGGANA 1512
QY 511 ASPLeuSerLeuTrpProLALTYRANLEUThRGLUGINTYRLEUGLINTYASPLEUAPR 530

```

```

DB 1513 GGCTTGCCACTGTGTGCCGCGCTTACAGTCAAGCCGAGCACTGAAGCTGATTTGAAC 1572
QY 531 MetSerLeuGLINARGLEUPLYSGIUPROARGVALASPPHEThrSerThrILEPRO 550
DB 1573 ATPAGCGTGGACAGAAACTGAAAGAGCAGAGTGTGAGTTTGTGTCAATCTCTCCC 1632
QY 551 LeuILEUSERLALASERAPMETLEUHIISERPROLEUSERSerLeuThrPHELEUSER 570
DB 1633 CTGATTAATGTCCATGTCCACAGCAACCCCTGTGCTCTCGTGTCCCTTACTCTCCCTTCT 1692
QY 571 LeuLeuGLINPROPHEPHEPHEPHECYALAPRO 581
DB 1693 GTGCTCTGCTTCTTCTTAATCTCTGTCTCTCT 1725

```

Search completed: June 15, 2005, 12:01:59
Job time : 7272 secs

this Page Blank (uspto)

Query	Match	Length	ID	Description
1	1282.5	41.7	2687	AK077248 Mus muscu
2	1272.5	41.3	2661	AK033563 Mus muscu
3	1264	41.1	1680	AY414461 Mus muscu
4	1236	40.1	1958	AK078951 Mus muscu
5	1230	39.9	1680	AY414459 Mus muscu
6	1197	38.9	4937	AK040349 Mus muscu
7	1191	38.7	2868	HS080313 Homo sapi
8	1182	38.4	3909	AL713761 Homo sapi
9	1158.5	37.6	2038	EX580806 Homo sapi
				BC019926 Mus muscu

10	1142.5	37.1	2356	3	AK037191	Nus muscu
11	1129.5	36.7	1444	3	AK078879	Nus muscu
12	1097.5	35.6	1818	7	CF111083	Shultzomi
13	1069.5	34.7	1826	3	CR618303	full_Leng
14	1055	34.3	1859	3	CR641368	Tetradodon
15	1048	34.0	1817	3	CR634885	Tetradodon
16	1038	33.7	1449	3	CR647187	Tetradodon
17	1028	33.4	1823	3	CR640788	Tetradodon
18	1026	33.3	1853	3	CR723405	Tetradodon
19	1020	33.1	1857	3	CR637401	Tetradodon
20	895.5	29.1	742	4	B1822069	AK050444
21	815	26.5	1610	3	AK050444	Nus muscu
22	804	26.1	1438	7	CF111142	Shultzomi
23	804	26.1	1438	7	CF111142	Shultzomi
24	780	25.3	4319	3	BC069040	Nus muscu
25	779	25.3	3853	3	AK039018	Nus muscu
26	768	24.9	2547	9	AY402702	Homo sapi
27	763	24.8	2173	3	AK046080	Nus muscu
28	763	24.8	2181	3	AK043748	Nus muscu
29	763	24.8	2184	3	AK042335	Nus muscu
30	760.5	24.7	2187	3	BC026315	Homo sapi
31	760	24.7	2026	3	AK052200	Nus muscu
32	755.5	24.5	1723	9	AY407736	Homo sapi
33	755.5	24.5	1723	9	AY407737	Pan trogl
34	754	24.5	1723	9	AY407738	Nus muscu
35	752	24.4	1810	3	CR611139	full_Leng
36	752	24.4	2177	3	AK036559	Nus muscu
37	752	24.4	2523	6	CD014089	AK036259
38	750.5	24.4	576	4	B1824830	603033728
39	746	24.2	3206	3	AK05037	Nus muscu
40	737	23.9	2293	9	AY402703	Pan trogl
41	733.5	23.8	1113	7	BC066114	F05_Ag2_P
42	726	23.6	1145	3	BC036813	Homo sapi
43	718.5	23.3	1117	7	BC061747	Ag2_P32_P1
44	717	23.3	818	7	CN067322	Nus_Ag2_P
45	716	23.3	3431	3	BC041823	Homo sapi

ALIGNMENTS

RESULT 1	AK077248	LOCUS	AK077248	2687 bp	mRNA	linear	HTC 03-APR-2004
DEFINITION							Mus musculus 11 days pregnant adult female ovary and uterus cDNA,

product: similar to CARBOXYL ESTERASE PRECURSOR (EC 3.1.1.1)
(AL-ESTERASE) (B-ESTERASE) (MONOESTERASE) (COCAINE ESTERASE)
(PROCAINE ESTERASE) (METHYLBUTYRASE) [Mesocricetus auratus], full
insert sequence.

ACCESSION	AK077248
VERSION	AK077248.1
KEYWORDS	GI:26346111
SOURCE	HTC; Cap trapper.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Mus

AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE	99279253
PUBMED	10349636
REFERENCE	2
AUTHORS	Carninci, P., Shibata, Y., Hayatsu

Ichih, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE	20499374
PMED	11042159
REFERENCE	
AUTHORS	3
	Shibata, K., Itch, M., Aizawa, K., Nagata, S., Sasaki, N., Carninci, P.,
	Kono, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itch, M.,

Db 1438 ATMAAACTT-----CCTTACACTGAGAGAGCACTAAGCAGAGATGATGAG 1491
 QY 497 TTTTTPAATThPhellaarThrglyAaenProaenGlyAaenAaenSerleuTPro 516
 Db 1492 TACCGGACCACTTTGACAGGCGACCGGAGCCCAACAGAGAGGCTTACCTTACCTGAGCTT 1551
 QY 517 AATATyAaenLeuThrglyGlnTyrLeuGlnLeuAaenAaenSerleuGlnGln 536
 Db 1552 GTGATGAGCAGCAGATGAGAGTACCTGAGCTGAGCATCAGACCTTCTGTGGGTGAGCC 1611
 QY 537 LeuLyGlnProAlyAaenAaenPheTyrPheSerThrllePro-----Leuileu 553
 Db 1612 CTGAAGGCCAGAGAGCTGAGTTCGAGCAAGACTCTGCCAGAGATCCAGAGACT 1671
 QY 554 SerAlaSerAaenPheLeuHisSerPro----- 562
 Db 1672 AAGGGGTCTCAGAGAGCAGCAAGAGTTCAGTTCCTGTGTGAAGAGAGCATGTGG 1731
 QY 563 -----LeuSerSerleuThrPheLeuSerleuLeuGlnProPhe 575
 Db 1732 CTTCATGCTGAGGTTTGGAGAGCTTTTGACCTTACCTTACCAACCACTTC 1785
 RESULT 3
 AY414461 1680 bp DNA linear GSS 17-DEC-2003
 LOCUS Mus musculus CES2 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY414461 GI:39770423
 VERSION AY414461.1
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1680)
 AUTHORS Clark,A.G., Gnanoweki,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanendau,D.M., Civeleo,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source 1..1680
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>1680
 /gene="CES2"
 /locus_tag="HCM5231"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,136-124 Length: 1680
 Score: 1264.00 Matches: 264
 Percent Similarity: 61.96% Conservative: 83
 Best Local Similarity: 47.14% Mismatches: 181
 Query Match: 41.05% Indels: 32
 Db: 9 Gaps: 10
 US-10-023-515-2 (1-581) x AY414461 (1-1680)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTyrCysePhePheLeuLeuGlnPro 20
 Db 16 CTTCCTGATGAGCTGAGTGCCTGCGCC-----TGGGGCTCTGGTT----- 57
 QY 21 LeuLeuGlyHisArgGlnTyrGlyThrGlyProSerAlaGlnGlyProGlnArgan 40
 Db 58 CTCCCTCCCAATGATAG-----GGCCAGAGCATCAACAGTCCCTCAGAGACC 105
 QY 41 ThrArgLeuGlyTyrPheGlnGlyLeuGlnValThrValLeuGlySerProValProval 60
 Db 106 ACACACACAGAGCCAGATCTGGGACGCTTATCCATATAGAGACCTTGATGGGTGTC 165
 QY 61 AenValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuAaenPheThrAaen 80
 Db 166 CACTCTTCTCTGGAAATCCCTTTGCCAAGCTCTGTAGAGACCATCCGCGCATGTGT 225
 QY 81 ProGlnProAlaSerProTyrAaenLeuAaenArgGlnAlaThrSerTyrProAaenLeuCyse 100
 Db 226 CCGAGGCCACCTGAACCAATGGGGTGTGTGAGATGGAGACATCACTCCGCGCATGTGT 285
 QY 101 LeuGlnAaenSerGlnTyrPheLeuLeuAaenPgin---HisMetLeuValHisTyrPro 119
 Db 286 CTGACAGACATCATCTGCATATGATAGAGAGCTTTAAGCTGTGAAGCTGACCTTGCTT 345
 QY 120 LysPheGlyValSerGlyAaenCyseLeuTyrLeuAaenIleTyrAlaProAlaHisAlaasp 139
 Db 346 CCTTCCCATGTCTGAGAGAGCTGCTTATACCTTAACCTAATGACCAAGATCATGCCAT 405
 QY 140 ThrGlySerLysLeuProValLeuValTyrPheProGlyGlyAlaPheLysThrGlySer 159
 Db 406 GAGGGTCCACCTTCCTGTGATGTGTGATGCCAGGGTGTGACCTGGTTCAGAGCATG 465
 QY 160 AlaSerIlePheAaenPginSerAlaLeuAlaAlaTyrGlnAaenValLeuValValVal 179
 Db 466 GCTTCATGATATGATGATCCGTCTGACAGCTAGAGATGTGTGTCTGCTATC 525
 QY 180 GlnTyrArgLeuGlyIlePheGlyPhePheThrTyrTyrAaenGlnHisAlaProGlyAaen 199
 Db 526 CAGTACCCCTGGGTGTCTGTGGTTCTTCAGACATGGGGACACAGATCCAGAGGCAT 585
 QY 200 TTPAlaPheLysAaenPginValAlaAlaLeuSerTyrValGlnLysAaenIleGlnPhePhe 219
 Db 586 TGGGATACCTGAGACCAAGTGGCTCCCTACGCTGGTCCAGCAAAATATCCCCCACTT 645
 QY 220 GlyGlyAaenProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSer 239
 Db 646 GTGGCAACCGTACCGGGGTATCAATTTTGGCGAGTCTGCGAGGCGACCAAGTGTGCTT 705
 QY 240 SerLeuIleLeuSerProMetAlaLysGlyLeuPheHisLysAlaIleMetGlnSerGly 259
 Db 706 TCACTGTGTGTCTCCCATGTCCTCAAGGGCTCTTCATGTGGCCATATAGAGAGTGG 765
 QY 260 ValAlaIleIleProTyrLeuGlnAlaHisAaenTyrGlnLysSerGlnAaenLeuGlnVal 279
 Db 766 GTGGCCCTGTGCTTACCTTATCTCAACACCTGAGATGTCTCC-----ACTAG 819
 QY 280 ValAlaHisPheCyseGlyAaenAaenAlaSerAaenSerGlnAlaLeuLeuAaenCyseLeuArg 299
 Db 820 GTGGCCAGCTCTCTGATGTGAGGCTGAGACTGAGAGCCCTGTGCGCTGTGTGAGA 879
 QY 300 ThrLysProSerLysGlnLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 319
 Db 880 GGCAAGAGTGAAGCAGAGATTTGTGCTATTAACAAAGACTTCCAGATATCCCTGTGTG 939
 QY 320 ValAaenGlyAlaPhePheProAaenGlnProLeuAaenPheLeuSerGlnLysAlaPheLys 339
 Db 940 GTGATGAGAGATTCCTACCAAGCATCTCAGAGAGCTGTGGCTCTGTGATTTTCAAC 999
 QY 340 AlaIleProSerIleIleGlyValAaenAaenHisGlnCyseGlyPheLeuLeuProMetLys 359
 Db 1000 CCTGTCCCAAGCATCATTTGTTTCAACATGATGATGTGTGATGTGTCCCAAAAGC 1059
 QY 360 GlnAlaProGlnIleLeuSerGlySerAaenLysSerLeuAlaLeuHisLeuIleGlnAaen 379

```

Db      1060 ATGAGC-----ATGATTGACTCCAGAGAGGAGTGGACAGACAAATGCTGCAAGCT 1110
Oy      380 ILeuHis-----1LeProGlnTyrLeuHisLeuValAla 392
Db      1111 ATCTTGACAGAGAAAGGCAACCCAGATGATGCTGCTCTTAAGTAGTACGCTGTAATG 1170
Oy      393 AaGlnTyrPheHisAaPlyHisSerLeuThrGlnLeuArgSerLeuLeuAaPlyLeu 412
Db      1171 GAAGAGTACATGAGGGGACCTAGAGGATGCCAGCCCTCCAAATACAGTTACAGAGAAATG 1230
Oy      413 LeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisAaPlyAla 432
Db      1231 ATGGAGATTTCATGTTTGTATGATCCCTGACCTCAAGTACATATTT---CAGAGTCC 1287
Oy      433 GlyAlaProValTyrPheTyrGluPheArgHisArgProGlnGlySerHisGluAspThrLys 452
Db      1288 CAGCCCTCTCTACTTCTTATGAAATTCACATCAATCCAGCTTCTCTCAAGAGATGCAG 1347
Oy      453 ProAlaPheValIleValAaPheHisAaPlyValArgPheValPheGlyGlyAlaPhe 472
Db      1348 CCACCCACAGTGAACCTGACATGAGATGAGGTTCTTTGTTTGTCTTCTTCTTC 1407
Oy      473 -----LeuIleGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeu 490
Db      1408 TGGGGCATAACTTACTTC-----ACTGAGGAGAGAGAGCTGCTG 1449
Oy      491 SerArgLysMetMetLysTyrTyrAlaThrPheAlaArgThrGlyAspProAsnGlyAsn 510
Db      1450 AACAGAGAGATGATGAAAGTACGCGGCACCTTGCACGACATGGAAACCCCAACAGTGA 1509
Oy      511 AaPlySerLeuTyrProAlaTyrPheLeuThrGlnGluGlnTyrLeuGlnLeuAspLeuAsn 530
Db      1510 GGCTTACCTACTGCTGTGTGATGAGCCATGATGAGAGTACCTGACAGTGCATCCAG 1569
Oy      531 MetSerLeuGlyGlyArgLeuLysGluProArgValAaPhePheThrSerThrIlePro 550
Db      1570 CTTTCTGCGGTGCGAGCCCTGAAGGCCAGAGAGGCTGACGTTCTGGACCAAGACTCTGCC 1629

```

RESULT 4
AK078953
LOCUS
DEFINITION
AK078953 1958 bp mRNA linear HTC 03-APR-2004
Mus musculus adult male cecum CDNA, RIKEN full-length enriched library, clone:9130231C15 product:similar to LIVER CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus], full insert sequence.
AK078953.1 GI:26347654
HTC; CAP trapper.
Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
1 Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
METH. ENZYMOLOGY 303, 19-44 (1999)
MEDLINE
PUBMED
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
GENOME RES. 10 (10), 1617-1630 (2000)
MEDLINE
PUBMED
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, T., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsuno, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
PUBMED
20530913
11076861
4
THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
(bases 1 to 1958)
Adachi, J., Aizawa, K., Akiyama, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Kahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 1958
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:9130231C15"
/db_xref="taxon:10090"
/clone="9130231C15"
/sex="male"
/tissue_type="cecum"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
66. 1742
/note="unnamed protein product; putative stimulant to LIVER CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus] (SWISSPROT|064419, evidence: FASTV, 78.4kD, 100kDa length, match=1674)"
/codon_start=1
/protein_id="BAC37476.1"
/db_xref="GI:26347655"
/translation="MPLRLPGMLCVVACGILLILQHVHGDSASPIRNTRGVGRGS FVHVYDTSVGHAFGIPFAPKPVGLRFAPDEPSPSGVSDGTSGPAMCLQDINYL LEAKENMLIPRIPMSDCLYINLYPTPAQDSNP.PVWWIHGGGLYASASAMNY DRAITETIVAIQVRLGVTGSPSTDOHARQKMGVGLDYVAALRNQKATAYFGNR SKVITFGVAGCTGVSHIISPMKGLFPHGAIKQSGVALPDLISDTSEVVYKTVANL SCGANTSEALIHCLRAKSKOETLIATINQVFKMTPAVVVGSEFLPKHPELITSMDFHIV

PSIGVNTDCEGMGVPMFMGLDHIKNITRETLPAVLKNYARMMPLPECSHLLIVEEY
 MGDDEPDLLOAFREMLGDPMFVPIPAIOVAHPFAGOPVRYEYEOHLSFKHYRPS
 HYKADHGDVAAPVPSYLMNDMLDLEEBELKRMVMYANFANGNPSGLPSWP
 VLDDHBOYLOQDPTOPAVGRALKAKRLOQMTKTLPKQIOELKSOQKAEAL"

polyA_signal

1937..1942

polyA_site

1958

/note="putative"

ORIGIN

Alignment Scores:

Pred. No.:	1,516-121	Length:	1958
Score:	1236.00	Matches:	261
Percent Similarity:	58.97%	Conservative:	81
Best Local Similarity:	45.00%	Mismatches:	187
Query Match:	40.14%	Indels:	51
DB:	3	Gaps:	8

US-10-023-515-2 (1-581) x AK078953 (1-1958)

QY 31 GYProSerAlaGluGlyProGlnIaArgAenThrArgLeuGlyTrrPrlGlnGlyLeuGln 50
 DB 141 GGGCCAGAGCTCAGCCAGCCCATCCAGAACCCAGGGGTCAGGTGAGAGGCGCTTT 200
 QY ValThrValLeuGlySerProValProValAenValPheLeuGlyValProPheAla 70
 DB 201 GTCCACGTGAAGACATTAATCTGTCTCCATGCTTCCTGGAAATCCCTTTTCCAG 260
 QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrrAsnLeu 90
 DB 261 CCTCTGTAGAGCTACGTGGCTTTGCACTCTGAGAACCTGAGCACTGAGTGTGTG 320
 QY 91 ArgGluAlaThrSerTrrProAsnLeuGlyLeuGlnIaAsnSerGluTrrLeuLeuAsp 110
 DB 321 AGAGATGAGACCTCAGCCGCGCATGTCTGTGCAAGCTGATTAATGATTAAGAGAT 380
 QY 111 GlnHisMetLeuValHisTrrProLysPheGlyValSerGluAspCysLeuTrrLeu 130
 DB 381 GCGAAGAAATGATCTGATCTGCTCCCATCTCTATGTCGAGAGCTGCTGTATCTC 440
 QY 131 AsnIleTrrAlaProAlaHisAlaAspThrGlySerTrrLeuProValLeuValTrrPhe 150
 DB 441 AACATCTACACCAACCAACGCGGCTTAACCTGCTGTGATGTATGATGATC 500
 QY 151 ProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 170
 DB 501 CACGGTGTGTGACGTGTGTGATGAGATCAGCTTCATGAATGATGATTAATGATGAC 560
 QY 171 TrrGluAspValLeuValValValGlnTrrArgLeuGlyIlePheGlyPheThr 190
 DB 561 ACTGAGAAATGATTAATGATCTGCTCCCATCTGATGCTGGGCTCTTGGCTTCTTAC 620
 QY 191 ThrTrrAspGlnHisAlaProGlyAsnTrrAlaPheLysAspGlnValAlaLeuSer 210
 DB 621 ACTGAGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 680
 QY 211 TrrValGlnLysAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGly 230
 DB 681 TGGGTGCGAAGAAATGATCTTATCTTGTGAGCAACCGTGAACAGGCTCACTATTTGGA 740
 QY 231 GluSerAlaGlyValIleSerValSerSerLeuIleLeuSerProMetAlaLeuGlyLeu 250
 DB 741 GTCTAGCAGGAGGTAACAAGTGTCTCTCAATATCTGCTCCCAATGTCACCAAGACTC 800
 QY 251 PheHisLysAlaIleMetGlnSerGlyValAlaIleIleProTrrLysGlnAlaHisAsp 270
 DB 801 TTTTATGAGGACCATTAATGAGAGTGAAGTGGCTGCTGCTGACCTATCTCT----- 854
 QY 271 TrrGlyLysSerGluAspLeu---GlnValValAlaHisPheCysGlyAsnAsnAlaSer 289
 DB 855 ---GACACTTCTGAGGTGTGATTAACAACAAGTGAACCTATCTGCTGTGAGAACCA 911
 QY 290 AspSerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGlnLeuLeuThrLeu 309

DB 912 GACTCAGAACCCCTATATCAGCTGCTGAGAGCAAGAGTAAACAAGATCTGCTGATTT 971
 QY 310 SerGlnLysTrrLysSerPheThrArgValValAspGlyAlaPhePheProAsnLysPro 329
 DB 972 AACCCAGTCTTCCAAATGATCCCTGCTGTGTGGATGTGGAGATTCTTACCCAAAGATCTT 1031
 QY 330 LeuAspLeuLeuSerGlnLysAlaPheLysValIleProSerIleIleGlyValAsnAsn 349
 DB 1032 CAGAGCTGCTGACCTCTATGATTTTCAACCTGCTGCTCCAGATCATTTGTCTCAACA 1091
 QY 350 HisGlyCysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsn 369
 DB 1092 GATGAATGTGGCTGGAGGATCCCATG-----TTCATGGGCTTTGAC 1133
 QY 370 LysSerLeuAlaLeuHisLeuIleGlnAsnIle----- 380
 DB 1134 -----CATATCATTAAGAACATTAACCAAGAGACCTGCTGCTTCTG 1178
 QY 381 -----LeuHisIleProProGlnTrrLysHisLeuValAlaAsnGlu 394
 DB 1179 AAGAACACAGACACGAGATGATGCTGCTCCAAAGATGATGATCCTGCTGATGAGAG 1238
 QY 395 TrrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeuAspLeuGly 414
 DB 1239 TATATGGGGGACATGAGAGACCCAGAACCTGCAAGACAGATTCAAGAGATGCTGGGG 1298
 QY 415 AspValPhePheValValProAlaLeuIleThrAlaArgTrrHisArgAspAlaGlyAla 434
 DB 1299 GACTTCATGTTGTGATGATCCCTGCTCCAAAGTGCACATTTCCAGAGCT---TCCAGAGCT 1355
 QY 435 ProValTrrPheThrGluPheAlaHisArgProGlnCysPheGluAspThrLysProAla 454
 DB 1356 CCTGTACTCTTATGAGTTCAGAGTCTGTCAGCTTCACTCAACAGATGTCAGGCACTCA 1415
 QY 455 PheValLysValAspHisAlaAspGluValArgPheValPheGlyGlyAlaPheLeuLys 474
 DB 1416 CATGGAAGGACAGACCATGATGATGACCTTGCTTGTGGCTGCTATTTGTGGAGC 1475
 QY 475 GlyAspIleValMetPheGluGlyAlaThrGluGlnLysValLeuLeuSerArgLysMet 494
 DB 1476 ATGAATTTGGACCTC-----ACTGAGGAGGAGAGCTGCTTAAGAGATGATGAT 1523
 QY 495 MetLysTrrTrrAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeu 514
 DB 1524 ATGAAGTACTGGCCAAATTTGCAAGAAATGGAATCCCAACAGTGAAGGCTTACCTTCC 1583
 QY 515 TrrProAlaTrrAsnLeuThrGlnTrrLysGlnLeuAspLeuAsnMetSerLeuGly 534
 DB 1584 TGGCTGTGCTTGGACCATGATGAGACATGACGTCGACGTCGACACCCAGCTGCTGGGT 1643
 QY 535 GlnArgLeuLysGluProArgValAspPheTrrPheThrIleProLeuIleLeuSer 554
 DB 1644 CGAGCCCTGAAGGCGCAAGAGGCTGCAATTTGCAACCAAGATCTGCCCAAGAAATCCAG 1703
 QY 555 -AlaSerAspMetLeuHisSerProLeuSerSerLeuThr----- 567
 DB 1704 GAGCTCAAGGTTCTCAGACACAAGCATCAGAGCTGTAGCCTGTCTGTCAAGGAAGCTG 1763
 QY 568 -----PheLeuSerLeuGlnProPhePhePhe 578
 DB 1764 TGTGAGATTTAAAGACATGCTGATTTCTGAGTTTACAAGATCATTTATTCATTTT 1821

RESULT 5

AY144459

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AY144459 1680 bp DNA linear GSS 17-DEC-2003

LOCUS Homo sapiens CES2 gene, VIRTU1 TRANSCRIPT, partial sequence.

DEFINITION genomic survey sequence.

ACCESSION AY144459

VERSION AY144459.1 GI:39770421

KEYWORDS GSS.

SOURCE Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 1680)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A., Todd, M.A., Tanenbaum, D.M., Civejlo, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J., Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
Science 302 (5652), 1960-1963 (2003)
14673302
2 (bases 1 to 1680)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A., Todd, M.A., Tanenbaum, D.M., Civejlo, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J., Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment.
Location/Qualifiers
source 1. 1680
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1. >1680
/gene="CES2"
/locus_tag="HCM5231"
ORIGIN
Alignment Scores:
Pred. No.: 5,25e-121 Length: 1680
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 39.95% Indels: 24
DB: 9 Gaps: 8
US-10-023-515-2 (1-581) x AY14459 (1-1680)
QY 31 GlyProSerAlaGlyGlyProGlnArgAsnThrArgLeuGlyTyrPileGlnGlyValGln 50
DB 76 GGGCCAGACTCAGCCAGTCCCTCCGACACACACAGGGGAGAGCTCTGGAGTCTT 135
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 136 GTCCATGTGAAGGGGCGCAATGCCGGGGTCCAAACCTTCCTGGGAATCCATTGCCAAG 195
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeu 90
DB 196 CCACCTCTAGGTCGGCTGCGATTGCAACCCCTGAGCCCTGAAATCTTGAGTGTGTG 255
QY 91 ArgGlnAlaThrSerTyrProAsnLeuGlySerGlnAsn-----SerGln 105
DB 256 AGGATGGAACCAACCATCCGCGCATGTCTTACAGACCTCACCGAGTGGAGTCCAG 315
QY 106 TrpLeuLeuLeuAspGlnHisMetLeuValHisTyrProValPheGlyValSerGln 125
DB 316 TTTCTT-----AGCCAGTTCACATGACCTTCCTCCGATCCGATGTGTAG 363
QY 126 AspCysLeuTyrLeuAsnIleTyrAlaProAlaHisAlaSerThrArgIleSerPro 145
DB 364 GACTGCTGTACCTCAGACATCTTACACCCCGGCGCATAGCAATGAAGGCTTAACTGCCG 423
QY 146 ValLeuValTyrPheProGlyGlyValAlaPheLeuThrGlySerAlaSerIlePheAspGly 165
DB 424 GTATGTGTGTGATCCACGGGTGTGCGCTGTGTTTGGCATGTGCTTCTGTATGTGTGT 483
QY 166 SerAlaLeuAlaIleTyrGlnAspValLeuValValValGlnTyrArgLeuGlyIle 185
DB 484 TCCATGTGTGCTGCTTGGAGAACGTGTGTGTGTATCATCAATCAACCGCTTGGGTGTTC 543

QY 186 PheGlyPhePheThrThrTyrAspGlnHisAlaProGlyValAsnTyrAlaPheLeuAspGln 205
DB 544 CTGGGCTTTCAGACACTGAGACAGACAGACGACACCGGCACTGGGGCTTACCTGACCA 603
QY 206 ValAlaAlaLeuSerTyrValGlnIleAsnIleGlnPheGlyGlyAspProSerSer 225
DB 604 GTGGGTGACATACGCTGGGTCCAGAGATATCCGCCACTTTGGAGGCAACCTGACCGT 663
QY 226 ValThrIlePheGlyGlySerAlaGlyValAlaIleSerValSerSerIleLeuSerPro 245
DB 664 GTCCACCATTTTGGGAGATCTGGGGTGGCACAGATGTGTCTTGGCTGTGTGTCTCC 723
QY 246 MetAlaIleGlyLeuPheIleValAlaIleMetGlySerGlyValAlaIleIleProTyr 265
DB 724 ATATCCCAAGGACTCTTCACAGAGCATCATGAGAGTGGCGTGGCTCTCCGCCG 783
QY 266 LeuGlnAlaHisAspTyrGlyIleValSerGlyLeuPhe-----GlnValValAlaHisPhe 283
DB 784 CTCATTGCC-----AGCTCAGCTGATGTATCTCCACGGTGTGGCCAACTCG 831
QY 284 CysGlyAsnAsnAlaSerAspSerGlyAlaIleuLeuArgCysLeuArgThrIleProSer 303
DB 832 TCTGCTGTGTACCAAGTTGACTGAGGCTGTGGGTGCTCTCCGGGCGCAAGTAA 891
QY 304 LysGlnLeuLeuThrLeuSerGlnIleValThrLysSerPheThrArgValValAspGlyAla 323
DB 892 GAGAGATCTCTTGCAATTACAAACCTTTCAGATGATATCCCGAGAGTGGGTGGGGT 951
QY 324 PhePheProAsnGlnProLeuAspLeuLeuSerGlnIleValHisIleIleProSer 343
DB 952 TTCCTGCCAGGACCCCGGAGAGCTGTGCTCTCCGACCTTTCAGCTGTCTTAC 1011
QY 344 IleIleGlyValHisAsnHisGlyCysGlyPheLeuLeuPro-----MetLys 359
DB 1012 ATGTGTGTGTCAACAATGAATTCGGGTGCTCATCCCAAGGATCATGAGATCTAT 1071
QY 360 GlnAlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaIleuHisIleuIleGlnAsn 379
DB 1072 GATTCACCAAGAAATGACAGAGAGGCTCCAGGCTGTCTTGAGAAATGTTTACG 1131
QY 380 IleLeuHisIleProProGlnTyrLeuHisIleValAlaAsnGlnTyrPheHisAspLys 399
DB 1132 CTGCTGATGTGCTCTCTCATTTGTGTACCTGTGAGGAGAGATCATTTGGGACAAAT 1191
QY 400 HisSerLeuThrGlnIleArgSerSerLeuLeuAspLeuGlyValAspValPhePheVal 419
DB 1192 GGGGATCCCAAGCCTCCAGAGCGAGTTCACAGAGATGATGCGGACATCCATGTTGTG 1251
QY 420 ValProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
DB 1252 ATCCCTGACCTCCAAAGTACACATTT---CAGTGTCCCGGGCCCTGTGTACTTCTAC 1308
QY 440 GlnPheArgHisArgProGlnCysPheGlnAspThrIleValProAlaPheValIleAsp 459
DB 1309 GAGTTCACATCAACCCGCTGCTCAAGAACATCAAGGCCACCGCACATGAAGCAAGC 1368
QY 460 HisAlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMet 479
DB 1369 CATGTGATGAGCTCTTGTGTTC---AGAAATTTCTTTGGGGGCAACTTACATAA 1425
QY 480 PheGlnGlyValThrGlnGlnIleValLeuLeuSerArgTyrMetMetLeuTyrTyrAla 499
DB 1426 TTC-----ACTGAGGAAGAAGAGAGAGCTTACAGGAAGATGATGATATCTGGGCC 1476
QY 500 ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTyrProAlaTyrAsn 519
DB 1477 AACTTGGAGAAATGGAAACCCCAATGGCGAGGAGTCTGCACACTGCGCTGTTCGAC 1536
QY 530 LeuThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLeuGln 539
DB 1537 CAGAGAGAGCAATACCTGACACTTACAGCTGTGGGTGGCCGGGCTTGAAGGCC 1596
QY 540 ProArgValaAspPheThrIlePro 550

Db	1597	CACAGGCTCAGTCTTGGAGAAAGGCGCTGCC	1629
RESULT 6		: :	
AK040349			
LOCUS			
DEFINITION	AK040349	4927 bp	mRNA linear HTC 03-APR-2004
			Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A330088B12 product:similar to CARBOXYESTERASE PRPCUSOR (EC 3.1.1.1) (AL-ESTERASE) (B-ESTERASE) (MONOBUSTRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE) [Mesocricetus auratus], full insert sequence.
ACCESSION	AK040349		
VERSION	AK040349.1	GI:26087790	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS			
TITLE	Carninci, P. and Hayashizaki, Y		
JOURNAL	High-efficiency full-length cDNA cloning		
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	99279253		
REFERENCE	10349636		
AUTHORS			
TITLE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	20499374		
REFERENCE	11042159		
AUTHORS			
TITLE	3 Shiba, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, T., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Ikawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)		
PUBMED	20530913		
REFERENCE	11076861		
AUTHORS			
TITLE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
JOURNAL	Functional annotation of a full-length mouse cDNA collection		
REFERENCE	Nature 409, 685-690 (2001)		
AUTHORS			
TITLE	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
REFERENCE	Nature 420, 563-573 (2002)		
AUTHORS	6 (bases 1 to 4927)		
TITLE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (16-JUL-2001) Yoshitake Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),		

[illegible]

[illegible]

OY	525	UGLLEuAspLeuAnuMetSerLeuGlyGlnArgLeuLeuGluProArgValaAapheTr	545
Db	1885	GCACCTGACATTCACCCCTTCTGTGGCCATGCCCTTAAGCCAGAAAGGCGCCATTTCTG	
OY	545	pThnSerThrIlePro	550
Db	1645	GACCAAGACTCTTCTCC	1660

RESULT	7
HSM803013	
LOCUS	HSM803013
DEFINITION	Homo sapiens mRNA; cDNA DKFpZ34N0935 (from clone DKFpZ34N0935).
ACCESSION	U1713761
VERSION	U1713761.1 GI:19584506
KEYWORDS	HTC.

SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS

Eukaryota, Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2888)
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Weissenreuther, R.

Mewes, H. W., Well, B., Amlid, C., Osanger, A., Fobo, G., Han, M., and Wiemann, S.

COMMENT
CONSTRM
TITLE
JOURNAL
Submitted (22-SEP-2004) MRS, Ingolstaedter Landstr.1, D-85764
Neuherrberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

FEATURES

Resequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp434N0935) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: <http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp434N0935>. Further information about the clone and the sequencing project is available at <http://mipe.gsf.de/projects/cdna/>.

Location/Qualifiers

Abstract

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RSPD:DKFZp434N0935"
/db_xref="taxon:9606"
/clone="DKFZp434N0935"
/issue_type="testis"
/clone_id="434 (synonym: htes3). Vector pSport1; host
DH10b; sites NotI + SalI"
/dev_stage="adult"
/note="carboxylesterase 2 isoform 2"
1..2888
gene
CDS
954..2777
/gene="DKFZp434N0935"

```

```

/gene="DKFZp434N0935"
/codon_start=1
/product="hypothetical protein"
/protein_id="CD28531.1"

```

```

/db xref="G1:19584507"
/db xref="UniProt/Swiss-Prot:000748"
/translation="MTAOSRSPTPTPFPGSGOPIPIPCVROGRLKALIHCTDPG
OPLGEOORVRRORTSETSEPTLRHLRIHRSAYACGILLIYVROGQSDASPIRTHT
GOWLSLIVHGANGAGVOTFPIGIPAKYDGLPLAFAPREPPESGVRDGTTHPACL
DOLTADESEPLISOPMTPRSDMSMEDCLYSLITPASHHRSKSNIPVWVTHGALVNG
MALSVDGSGMTAIIENYVYVITIQRLGVGTFSTEDKRIATNGWLDIVVALRVVOONT
MARGENPDRPTITFGESAGTIVSVLSVPSIQGPHALIMESGALLPGLIASASVLT
SRYVYANLACQVDSBALVGLRKSKEELIANKPRKMPGVYDGLTFPIRHOECLA
SADPQSVESIVGNNNEFGMLIPVMYIYDQCKMEASQALOKMLTYLMIPIFG
DLARESIYIGDNGDPOTLQVAGQOEMWASMPVTPALQVHOCGRARAYTEFPHOQSM
LKNIRPPIKMDHVKTEEBEOLSRKMKWAKVWVWVWVWVWVWVWVWVWVWVWVWVW
LQNTQPAQGRALKHRLQFMKCALPOITOELEENNEENGTETL"

```

Urgent

Alignment Scores:
red. No.:

2.01e-116 Length: 2888

Score:	1191.00	Matches:	242
Percent Similarity:	62.15%	Conservative:	88
Best Local Similarity:	45.57%	Mismatches:	151
Query Match:	38.68%	Indels:	40
DB:	3	Gaps:	8

US-10-023-515-2 (1-581) x HSM803013 (1-2888)

QY	31	GIYProSerAlaGLUGLYProGlnArgAsnThrArgLeuGLYTrpLleGlnLYuArgIn	50
Db	1221	GGCCAGAGCTACGACCACTCCATCCGAGCCACACACAGGGGCAAGTCTGGGAGTCTT	1280
QY	51	ValThrValLeuGLYSerProValProValAsnValPheLeuGLYValProPheAlaIa	70
Db	1281	GTCCATGTAAGGGGGCCCAATGCGGGGTCCAAACCTTCCTGGGAATTCATTTGGCAAG	1340
QY	71	ProProLeuGLYSerLeuArgPheThrAsnProGlnProLaseProTrpAsnLeu	90
Db	1341	CCACCTCAAGATCCGCTCGATTGGACCCCTGAGCCCTCGAATCTTGACATGGTGTG	1400
QY	91	ArgGlnAlaThrSerTrpProAsnLeuCysLeuGlnAsn-----SerGln	105
Db	1401	AGGATGGACCAACCCATCCGGCCATGATGTCTACAGACCTCACCGCATGGATCAAG	1460
QY	106	TrpLeuLeuAspGlnIleMetLeuIlyValIleValIleValIleValIleValSerGln	125
Db	1461	TTTCTT-----AGCCAGTTCAACATGACCTTCCTTCGAGTCCATGCTAG	1500
QY	126	AspCysLeuTrpLeuAsnIleTrpAlaProAlaIleValAspThrGlySerIlyLeuPro	145
Db	1509	GACGCGCTGTACCTCACATCTACACGCGCGCCCATACCAATGAAGCTCAACCTGCG	1568
QY	146	ValLeuValTrpPheProGlyGLYValAlaPheIlyThrGlySerAlaSerIlePheAspIly	165
Db	1569	GTAAATGGTGAATCCACGAGTGGTCCCTGTTTGGCATAGGCTCTTGATATGATGT	1628
QY	166	SerAlaLeuAlaAlaTrpGLYAspValLeuValValValGlnTrpArgLeuGLYIle	185
Db	1629	TCCATGCTGCGCTGCTGGAGAACGTTGGTGGTCATCATCCAGTACCGCTGGGTGTC	1688
QY	186	PheGlyPhePheThrThrTrpAspGlnIleValProGlyAsnTrpAlaPheIlyAspGln	205
Db	1689	CTGGGCTTCTTCACACATGACACAGACCAACCGGCACTGGGGCTACCTGGACCA	1748
QY	206	ValAlaAlaLeuSerTrpValGlnIlyAsnIleGluPheGlyGLYAspProSerSer	225
Db	1749	GTGGCTGACTACGCTGGGTCCAGCAAGATATCCGCCACTTTGGAGGCAACCTGACCT	1808
QY	226	ValThrIlePheGlyGluSerAlaGLYAlaIleSerValSerSerLeuIleLeuSerPro	245
Db	1809	GTACACATTTTGGGAGATCGCGGGTGGACAGATGTGTCCTGGCTTGTGTGTCGCC	1868
QY	246	MetAlaIyGLYLeuPheHisIlyValIleMetCysLeuGLYValAlaIleIleProTrp	265
Db	1869	ATATCCCAAGACCTTTCACAGAGCCATCATGAGAGTGGCGTGGCCCTCGCCCGGC	1928
QY	266	LeuGlnAlaIleAspTrpGlnIlySerIlyAspLeu-----GlnValValAlaIlePhe	283
Db	1929	CTCATTTGCC-----AGCTCACTGATGTCATCTCCACGATGTGGCCAACTG	1978
QY	284	CysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuAspCysLeuArgThrIlyAspSer	303
Db	1977	TCTGCGGTGACCAAGTGACTGTGAGGCCCTCGGTGGCTGCTGCGGGCAAGATAA	2036
QY	304	LYeGlnLeuLeuThrLeuSerGlnIlyThrIlySerPheThrArgValAlaAspIlyAla	323
Db	2037	GAGAGATTTCTTGCAATTAAACAAGCTTTCAAGATGATCCCGAGTGTGATGGGGTTC	2096
QY	324	PhePheProAsnGlnProLeuAspLeuSerLeuGlnIlyValaPheIlyAlaIleProSer	343
Db	2097	TTCTCGGCCAAGCACCCCAAGAGCTGTGGGCTCTGCGCACTTTCACGCTGTCCCTAAG	2156
QY	344	IleIleGlyValaAsnAsnIleGlnCysGlyPheLeuLeuPro-----MetIly	359

[illegible]

```
source
1..3909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686H0466"
/db_xref="taxon:9606"
/clone="DKFZp686H0466"
/tissue_type="endometrium carcinoma cell line"
/clone_id="686 (synonym: hlcc3). Vector pSPORT1_Sfi1, host
DH10B, sites SfiIA + SfiIB"
/dev_stage="adult"
/notes="cardoxylesterase 2 isoform 2"
gene
1..3909
/gene="DKFZp686H0466"
1002..2825
/gene="DKFZp686H0466"
/db_codon_start=1
/product="hypothetical protein"
/protein_id="CAD98009.1"
/db_xref="GI:31874223"
/db_xref="UniProt/Swiss-Prot:O00748"
/translation="MTAQSRSPTTPPPPSQRPRTLPQVQTPRRKALHICWTDPG
OPLEGQVRVROKRTSEPPMHLRLRLSAVAGLLLVKGQSDASPTRTTHT
GOVLSLVVKKANNAGVQTPLGIPAKPLGLRFAPEPPPSWGGVDRGTTHATC
OULTVASERLSPNMTFSPDSMSBCLYLSITPAHSHGSLNPLVMTWIHGALVFG
MSLIDGSKMLALENVVVI IQRLVGLFSTGDHATNGSLDYVALRWQONI
AFHGDNPRTVTFEGSAGTSVSLVSVSI SQELFSGALIMESGVALPLGLIASDVI
SVVVALSACDQVDSALVGLGRKSKSEIILAIKPKFKIPGVVGVLPRLPBLIA
SADPOVPSIVGVNNHFGMLPKVMRIYDTQKEMDRBSAALQKMLTLMLPTSG
DLIREYIDGNDPOTLQAFQPMADSWFVPLQVAFROCSRAPIYTFERHOPSM
LNIRPPIKADHVTFEERQLSRKMTKMTANPARNKPNESGLPHMPLFDQEROYL
RNLQPAVARALKARLPQPKALPQKIQLEPEPERHTEL"
ORIGIN
Alignment Scores:
Pred. No.: 3,05e-115 Length: 3909
Score: 1182.00 Matches: 241
Percent Similarity: 61.96% Conservative: 88
Best Local Similarity: 45.39% Mismatches: 162
Query Match: 38.39% Indels: 40
Dbs: 3 Gaps: 8
US-10-023-515-2 (1-581) x HSN806270 (1-3909)
QY 31 GTPProseraIaGlUGLyProGlnArgAenThrArgLeuGlyTTPILeGlnGlyLVaGln 50
Db 1269 GGGCCAGACTCCGCCAGTCCATCCGACACACACAGCGGGCAGGTGCTGGGAGTCTT 1328
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
Db 1329 GTCCATGTGAAGGGGCCAATGCCGGGCTCCAAACCTTCTGGGAATTCCATTGCCAAG 1388
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeu 90
Db 1389 CCACCTCTAGGTCCGCTCGCATTTGCAACCCCTGAGCCCTCGATCTTGGAGTGGTGTG 1448
QY 91 ArgGlnIaThrSerTyProAsnLeuGlyLeuGlnAsn-----SerGln 105
Db 1449 AGGGATGGAACACCACTCCGCGCATGTGTACAGACCTCACCGCAGTGGAGTCCAGAG 1508
QY 106 TrpLeuLeuLeuAspGlnHisMetLeuLeuValHisTyProLyPheGlyValSerGln 125
Db 1509 TTTCTT-----AGCCAGTTCACATGACCTTCCTTCGATCGCATGTCTGAG 1556
QY 126 AspCyLeuTyTrpLeuAsnIleTyAlaProAlaHisAlaAspThrGlySerLybLeuPro 145
Db 1557 GACTGCTGTACTTCAGACATCTACACGCCGCCCATGACCATGAAGGCTTAACTTGCCGC 1616
QY 146 ValLeuValIleTrpPheProGlyGlyAlaPheLeuThrGlySerAlaSerIlePheAspGly 165
Db 1617 GTGATGGTGTGGATCCACGGTGGTGGCTTGTTTTGGCATGCTTCTGTATGATGGT 1676
QY 166 SerAlaLeuValAlaTyArgIleAspValLeuValValValGlnTyArgLeuGlyIle 185
```

```
Db 1677 TCCATGCTGCTGCTGCTTGGAGAACGTTGGTGTGATCATATCCATACCGCTGGTGTCTC 1736
QY 186 PheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTTPAlaPheLybAspGln 205
Db 1737 CTGGGGCTTCTTCAAGCTCGAGAACAGACCAACCGCAACTGGGGCTGATCTGGACCA 1796
QY 206 ValAlaAlaLeuSerTyTrpValGlnLybAsnIleGluPhePheGlyGlyAspProSerSer 225
Db 1797 GTGGCTGACATACCGTGGTCCAGCAGAAATTCGCCCACTTTGAGGACCACTGACCGCT 1856
QY 226 ValThrIlePheGlyGlyLeuSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPro 245
Db 1857 GTACACATTTTGGGAGTCTGCGGGTGCACGAGTGTGTCTTGCTGTGTGTCCCC 1916
QY 246 MetalAlaLyLeuPheHisLeuAlaIleMetGlySerGlyValAlaIleIleProTy 265
Db 1917 ATATCCACAGACCTTCTCCACGAGCCATCATGGAGATGGCGCTTCCGCCGCC 1976
QY 266 LeuGlnAlaHisAspTyArgIleLybSerGlnAspLeu-----GlnValValAlaHisPhe 283
Db 1977 CTCAATTGCC-----AGCTCAGATGATGTCATCTCCACCGGTGGTGGCAACTG 2024
QY 284 CysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCybLeuArgThrLybProSer 303
Db 2025 TCTGCCCTGTACCAAGTGTGACTGTGAGGCCCTGTGGTGGCTGCTGCGGGCAAGATMA 2084
QY 304 LybGlyLeuLeuThrLeuSerGlnLybThrLybSerPheThrArgValValaAspGlyAla 323
Db 2085 GAGAGAGATCTTGTCAATTAACAAGCTTTCAGATGATCCCGGAGTGGTGGATGGGCTC 2144
QY 324 PhePheProAsnGluProLeuAspLeuLeuLeuSerGlnLybAlaPheLybAlaIleProSer 343
Db 2145 TTCCTGCTCCAGGACCCCCACGAGCTGTGGCTCTGCGCAGATCTGCTGCTGCCCTAGC 2204
QY 344 IleIleGlyValaAsnAsnHisGlyCybGlyPheLeuLeuPro-----MetLyb 359
Db 2205 ATTGTGTGTGTCAACAATGAATGATTCGGTGTCTATCCCAAGTCAATGAGATCTTAT 2264
QY 360 GlnAlaProGlnIleLeuSerGlySerAsnLybSerLeuAlaLeuHisLeuIleGlnAsn 379
Db 2265 GATACCCAGAGAAATGACACAGAGAGGCCCTCCAGCTGCTCTCGAGAAATGTTPAACG 2324
QY 380 IleLeuHisIleProProGlnTyLeuHisIleuValAlaAsnGlnTyPheHisIleAspLyb 399
Db 2325 CTGCTGATGTGCTCTCTATCATCTGTGTGACCTGTGAGGAGAGATCACTTGGGACAT 2384
QY 400 HisSerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheVal 419
Db 2385 GGGGATCCCAAGACCTCTCCAGCGCAGTTCCAGAGATGATGCGCGGATCCATGTTTGTG 2444
QY 420 ValProAlaLeuIleThrAlaArgTyHisArgAspAlaGlyAlaProValTyPheTy 439
Db 2445 ATCCCTGACCTCCAAAGTACACATTTT--CAGTGTTCGGGGCCCTGTGTATCTTAC 2501
QY 440 GluPheArgHisArgProGlnCybPheGlnAspThrLybProAlaPheValLybAlaAsp 459
Db 2502 GAGTTCAGACATCAAGCCAGCTGGCTCAAGAACATCAGGCCACGSCAGATGAAGCAAC 2561
QY 460 HisAlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLybGlyAspIleValMet 479
Db 2562 CAT-----GTAAATTC----- 2573
QY 480 PheGlnGlyAlaThrGlnGlnGlnLybLeuLeuSerArgLybMetMetLybTyTrpAla 499
Db 2574 -----ACGAGAGAAAGAGAGACGATTAAGAGAAATGATGATGATCTGGGCC 2621
QY 500 ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTyProAlaTyAsn 519
Db 2622 AACTTTGGAGAAATGGAGAACCAATGGCCAGAGGTGTGCACACTGGCGCTGTGCAC 2681
QY 520 LeuThrGlnGlnTyTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLybGln 539
Db 2682 CAGAGAGACAAATCTGCGGCTGAACCTTACAGCTCGGGGTGGCGGCTTGAAAGGCC 2741
```

```

Oy      540 ProkaryoticProteinTherapeuticPro 550
Db      2742 CACAGGCTCCAGTCTGGAGAGAGCGCTGCC 2774

RESULT 9
BC019926
LOCUS      BC019926      2038 bp      mRNA      linear      HTC 20-SEP-2002
DEFINITION Mus musculus, clone IMAGE:5123923, mRNA.
ACCESSION  BC019926
VERSION     BC019926.1 GI:18044766
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2038)
Strausberg, R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey R. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgobcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hui, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNLT at: http://image.llnl.gov
Series: IRAX Plate: 42 Row: P Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein
This clone has the following problem: frame shifted.

FEATURES
Source
1..2038
Location/Qualifiers
1..2038
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5123923"
/tissue_type="Liver, normal, 5 month old male mouse."
/clone_id="NCI CGAP L19"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ALIGNMENT SCORES:
Pred. NO.: 3.7e-113 Length: 2038
Score: 1158.50 Matches: 255
Percent Similarity: 59.62% Conservative: 86
Best Local Similarity: 44.58% Mismatch: 183
Query Match: 37.63% Indels: 48
DB: 3 Gaps: 12

US-10-023-515-2 (1-581) x BC019926 (1-2038)
Oy      1 MetProteinTherapeuticProteinPro 18
Db      32 ATACAGCTGGG-----TCAGAGTCTCTAGCTGGAGACTGTCTCTCTGACATT 85
19 GlnProteinTherapeuticProteinPro 38

```

[illegible]

```

Db      1112 GGAAGACCTGTGGAGATTTCAGGCGCTTCTTGCC-----ATTAT 1153
Qy      381 UHAIIEIPROPGIINTYrleuHlSLeuValAlaAenGlyTyreHhIsaplyshIsse 401
Db      1154 GGAAGTGGCCCCCTGATCATGATCCACCTGCATCATGAAATACATAGACAATGCGTCAGA 1213
Qy      401 rleuThGclulIeArGaSPserleuAapleuApleuLeuGlyAAspValPhehValValPr 421
Db      1214 TCATATCGATACAGATGATGCTCCAGCAATTCCTGGAGATATCTCGTTCATCATCTCC 1273
Qy      421 cAlaLeuIleThAlAaRgTyriSaRgaSPAlaGlyAlaProValTyrePheTyrgluP 441
Db      1274 TACCTTGAACTTCCAAATATACCTTCGAGATGCTGGGTGCTGTTTCTTGACAGATT 1333
Qy      441 eArGhIsaRgProGlnCySPhegluAapThlySProAlaPheVallyeAlAaSPhIsAl 461
Db      1334 CCAAGCATACCCCAAGTCTTTTGCAAGTTCAAGCAGCTGGGTGAAGGCTGACCATGC 1393
Qy      461 aAaPgluValaRgPheValPheGlyAlaPheLeu-----LyAGlyAaSPlleValIe 479
Db      1394 GTCCTGAGAAATTCCTTGTGTTTGGAGGTCTCTTCATCATGATGAGAGTCCCTTGGC 1453
Qy      479 rPhegluGlyAlaThGluGluGlyLeuLeuSerArygLyMetMetLySTyTTPAl 499
Db      1454 CTTCACAGAGGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1513
Qy      499 aThrPheAlaArGThGlyAaSProAenGlyAaSPleuSerleuTyroProAlaTyrs 519
Db      1514 CCAAGTTTGACGACGACGAGAAACCCCAATGCAAGGAGGCTGCTCTTGGCCCAATTAA 1573
Qy      519 nleuThrGluGlnTyrleuGlnleuAaSPleuAenMetSerleuGlyAlaArGlyLeuG 539
Db      1574 CCAAGTTGAACAATCTTGAGATGCTGTATGATCAACGAGCTGGGGGTGAAGCTAAAGA 1633
Qy      539 uProArGValaAaSPheThrThSerThrIlePro 550
Db      1634 GCGTCGCGCTACAGTTCTGAGACAGAACACTGCCA 1667

```

RESULT 10
AK037191 2356 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 6 days neonate skin cDNA, RIKEN full-length enriched library, clone:A030007M20 product:hypothetical Carboxylesterases type-B containing protein, full insert sequence.
ACCESSION AK037191.1 GI:26085654
VERSION AK037191.1 GI:26085654
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Mech. Enzymol. 303, 19-44 (1999)
MEDLINE 99379253
PUBMED 10349636
AUTHORS 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
AUTHORS 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, K., Tanaka, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Izawa, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

```

TITLE
JOURNAL RIKEN Integrated sequence analysis (RISA) system--384-format
MEDLINE sequencing pipeline with 384 multiplexed sequencer
PUBMED Genome Res. 10 (11), 1757-1771 (2000)
AUTHORS 20530913
REFERENCE 11076861
AUTHORS 4
TITLE The RIKEN Genome Exploration Research Group Phase II Team and the
JOURNAL PANTOM Consortium.
REFERENCE Functional annotation of a full-length mouse cDNA collection
AUTHORS Nature 409, 685-690 (2001)
5
TITLE The PANTOM Consortium and the RIKEN Genome Exploration Research
JOURNAL Group Phase I & II Team.
REFERENCE Analysis of the mouse transcriptome based on functional annotation
AUTHORS of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
6 (bases 1 to 2356)
AUTHORS Adachi, J., Aizawa, K., Akinura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, Y., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
DIRECT SUBMISSION
COMMENT Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
location/qualifiers
1..2356
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM,DB:A030007M20"
/db_xref="Ensembl:10090"
/clone="A030007M20"
/issue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="6 days neonate"
1..2356
/note="Hypothetical Carboxylesterases type-B containing protein (InterPro|IPR002018, evidence: InterPro)"
ORIGIN
misc_feature
Alignment Scores:
Pred. No.: 2,486-111 Length: 2356
Score: 1142.50 Matches: 243
Percent Similarity: 60.18% Conservative: 100
Best Local Similarity: 42.63% Mismatches: 182
Query Match: 37.11% Indels: 46
Gaps: 10
DB: 3
US-10-023-515-2 (1-581) x AK037191 (1-2356)
Qy 16 leuIleuGlnProleuLeuGlyHhIsaRgIIntPrGlyTyThrGlyProSerAlaGlu 35
|||||

```

```

Db 133 CTGTAAGTCACAGACGCTTAGGT-----GCTTGGACACCAAG 171
Qy 36 G1yProglInaArgAsnThrArgLeuGlyTrpIleGlnGlyValInhrValLeuGly 55
Db 172 GAACCTCTCCATGATCAACAGGATCTCCAGGAAAGAAAGAGATGATG----- 225
Qy 56 SerProValProValAsnValPheLeuGlyValProPheAlaAlaProProLeuGlySer 75
Db 226 GGAGACACACCCATCAAGTCTTCTGGGGATCCCTTCTTAAACCTCCGGTGGGACAC 285
Qy 76 LeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArgGluAlaThrSer 95
Db 286 CGCAGATTGCTCTCCACAGACCCCGCTACCTGGAACCGCATCAGATGCCACACCC 345
Qy 96 TyrProAsnLeuCyLeuGlnInaAsnSerGluTrpLeuLeuLeuAspGlnHisMet---Leu 114
Db 346 TAACCACTTCTGCTGCTTCAAGAAATCTGGGGACAGATATATATCCATGATCACTTAACT 405
Qy 115 LysValHisTyrProLysPheGlyValSerGluAspCyLeuTyrLeuAsnIleTyrAla 134
Db 406 CGGAACACAGTAAATGCGCTGATTCAGTGAAGACTGCTCTACTGAATGTGATGCC 465
Qy 135 ProAlaHisAlaAspThrGlySerLysLeuProValLeuValTyrPheProGlyGlyAla 154
Db 466 CCAGTACTCGCCCTGGGGCCCTCTGCTGCGACAGTGAATGTTGGTTTCTTGAGGGTGC 525
Qy 155 PheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAlaTyrGluAspVal 174
Db 526 TTCTCGCGCGTTCGGCTTCCACCTACCAAGGCTCAGAGTTGGCGCCGCTGGGAAAGTG 585
Qy 175 LeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrTrpAspGln 194
Db 586 GTGCTGCTGTTTCTGACGTACAGCTGAGCTGGGACCTCGGGCTTTTACGACACCGGCAACGC 645
Qy 195 HisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLys 214
Db 646 CATGCCCGCGGGAATTTGGGGGCTGCTGCGACAGATAGCTTTCGCGTGGTGGACGAGG 705
Qy 215 AsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGluSerIleGly 234
Db 706 AACATCGAAGCCTTTGGTGAAGACCCGACAGTGAACCTATTTGGTCAAGTGGCGGGA 765
Qy 235 AlaIleSerValSerIleLeuSerProMetAlaLysGlyLeuPheHisIleVala 254
Db 766 GCGATGACGCTTCAAGACTGAATGATACCCCTTAAGCCAGGCTGTGTTCCATCAAGCC 825
Qy 255 IleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAspTrpGluLysSer 274
Db 826 ATCTCCAGAGTGTGATGCGGGTCTGAAAGCCTTCATCACTCCGATCCAGTCACTGACGTCA 885
Qy 275 GluAspLeuGlnValValAlaHisPheCySerGlyAsnAsnAlaSerAspSerGluAlaLeu 294
Db 886 -----GCCAAGACAGATCGTCACTGCGTGGCTGCGACCAATTAAGCAAGAGGTGATG 939
Qy 295 LeuArgCyLeuArgThrLysProSerLysGluLeuLeuThrLeuSerGlnLysThrLys 314
Db 940 GTGGAATGCTGAGGCTCTCTCAAGAGAGAGTGAACGCTTTTCCAGAGGATGTCA 999
Qy 315 SerPhe-----ThiArgVal 319
Db 1000 TTCTTTAGGGCTACTCCCATTAAGATCCAAAGAGATCGTGGTTCCTGAGCCCAAGTG 1059
Qy 320 ValArgGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLys 339
Db 1060 GTGATGTGTGTGTTCTCCAGAAATCTGTGGTCTTCTGACCCGTGGGAGGATTTAA 1119
Qy 340 AlaIleProSerIleIleGlyValAsnAsnHisGlu-----CysGlyPheLeu 355
Db 1120 CCTGTGCTTACTCTCTAGGTGTCAACACGAGATTCGATGGAATTAACCTTTTCTC 1179
Qy 356 LeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHis 375
Db 1180 ATGAAGATTACAGCTAAATCAAGGTATGATGAACCAACCAATCATCAAAATTAATCTG 1239

```

```

Qy 376 LeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisLeuValAlaAsnGluTyr 395
Db 1240 AGTTTCACAGTTCCTGTGAATATATCACTAAGACAGAGTCCGAGTGTGGCAGAGATAT 1299
Qy 396 PheHisAsp-----LysHisSerLeuThrGluIleArgAspSerLeuLeuAspLeu 412
Db 1300 TTGAATGACGCTATCATGATTAACCATGATCGAAGATGTTCCGAAATCGCTGTATGACCTTA 1359
Qy 413 LeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAspAla 432
Db 1360 ATCGAGAGATGCCACCTTTGTCTACAGCCACATCCCAAGCTGCGAGTACCAAGATGCT 1419
Qy 433 G1yAlaProValTyrPheTyrGluPheArgHisArgHisArgProGlnCysPheGluAspThrLys 452
Db 1420 GGTCTCCCGCTTATCTGTATGACGTTCAAGACCATGCTTCTTGGGACATCATCAAA 1479
Qy 453 ProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGlyAlaPhe 472
Db 1480 CCT---TCAATGAGACAGACCAAGAGAGAGCTTTCCTATCTCTTGGCAGCCCTTTC 1536
Qy 473 LeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuSerArg 492
Db 1537 TCCAAAGGC-----TCATCTGCAGGTGAGGAAAGGAATTCAGCCTTC 1578
Qy 493 LysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeu 512
Db 1579 CAGATGATGAATATCTGGGCCCACTTGTCTCACTCGGAAACCCAAATGCGAAAGCTG 1638
Qy 513 SerLeuTrpProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuMetSer 532
Db 1639 CCTACTGGCCACCGTTTGACAAGATGAAGATGACGAGATGTCATTCGATCAAGA 1698
Qy 533 LeuGlyGlnArgLeuLysGluProArgValAspPheTrpThrSerThrIleProLeuIle 552
Db 1699 GTGGGTGTGAAGCTCAAGAGAAAGAAAGTGGCTTTCGTGAG-TCAGCTTTCGCAATC--- 1754
Qy 553 LeuSerAlaSerAspMetLeuHisSerPro 562
Db 1755 -----TCAGAAACCTTAGAGACCA 1775

RESULT 11
AK078879 1944 bp mRNA linear HTC 03-APR-2004
LOCUS AK078879
DEFINITION Mus musculus adult male colon cDNA, RIKEN full-length enriched
library, clone:9030604P03 product:carboxylesterase 3, full insert
sequence.
ACCESSION AK078879
VERSION AK078879.1 GI:26347580
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL MEDLINE 99279253
PUBMED 10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subcloning of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL MEDLINE 20499374
PUBMED 11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
Komno, H., Akizawa, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

```


Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, T. RIKEN integrated sequence analysis (RISA) system -384-Format sequencing pipeline with 384 multichannel sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913
MEDLINE
11076861
PUBMED

4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6
(bases 1 to 1944)
Fukuda, S., Furuno, M., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, Y., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nishizaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Shinkawa, A., Shiraki, T., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akihara, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenhiro-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-research.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/.
URL: http://fantom.gsc.riken.jp/.
Location/Qualifiers

1. 1944
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM:DB:9030604P03"
/db_xref="taxon:10090"
/clone="9030604P03"
/sex="male"
/issue_type="colon"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
64. 1761
/note="unnamed protein product; carboxylesterase 3 (MGI:2148202, GB|NW_051200, evidence: BLASTN, 99%, match=1939)
putative"

CDS
/codon_start=1
/protein_id="BAC37439.1"
/db_xref="GI:26347581"
/translation="MGKYPILVSLAAGTACWGPSPPVNTYKGLKTYNLESEFT QPAAVAGVPAKPPGLSLKRPQPEPSPVFKNTSTYPMKSGQAVGQVISELFT NRKENIPLQSEDLVNTYTPADLTSLNPMVWHGGGLVVGASITVLSLASH ENNVVVTI QYRLI WGFSTGDSRNGMGLDQVALRVQONIINFGNGNPSTVILF GSGAGSVSLVLSPLAKNLPHRAISESVSLTALITTDVAPINGLVATLISGCTT

TSAMVHCLROKTEDELLESLKLNLFKDLILGNPKESYPLPTVIGDVLPAKPEET
IAEKSFSTVPYIVGINKQEFWIIPTLMGYPLAGKDKDKTANSLIMKSYPTLKISEN
MI PVAAEYKLGITDLDITKKDLFQDLMADVGVPSVYSRSHDAGASTYVFEFR
PSFVSAMPKAVIGIDGDEIFSVSGPDLKQGAEEETNSIKYMKFPANRAGNP
GGGPHMEYDQKGYLKIGASTQDAQRLDKQKVSFPAELRAKESAPRSRHEVEL"
1922. 1927
/note="putative"
polyA_signal
polyA_site
1944
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 4,61e-110 Length: 1944
Score: 1129.50 Matches: 249
Percent Similarity: 56.02% Conservative: 72
Best Local Similarity: 43.46% Mismatches: 205
Query Match: 36.68% Indels: 47
DB: 3 Gaps: 10
US-10-023-515-2 (1-581) x AK078879 (1-1944)
QY 16 Leu1leuGlnProleuLeuGlyNIaArgGlnTrpGlyLysThrGlyProSerAlaGlu 35
|||
DB 79 CTGATATGGCTTCTCTGCTGCGTGACAGCTGGGGGTAC-----CCATCCTACCA 132
QY 36 GlyProGlnArgThrPheLeuGlyLysThrGlyLysGlnValThrValLeuGly 55
|||
DB 133 ---CCTGTGTGAACACTGTTAAAGCAAGCTCTGGGAAGATGTCATTTGGAAAGGA 189
QY 56 SerProValProValaenValPheLeuGlyValProPheAlaAlaProProLeuGlySer 75
|||
DB 190 TTGACACAGCGCTGGGTGTTTCTCGGAGTCCCTTTCACAGCCCTCTTGCTCC 249
QY 76 LeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArgGluAlaThrSer 95
|||
DB 250 TTGAGATTTCCTCACACAGCGCTGAGAGCCCTGGAGCTTGTGAAGAACACCACTCC 309
QY 96 TyrProAsnLeuCyLeuGlnAsn-----SerGluTrpLeuLeu 108
|||
DB 310 TACCCGCTATGTCCTTCCAGATGCTGTGGAGGCGAGGCTCTTCAAGCTTTCACC 369
QY 109 LeuAspGlnHisMetLeuLysValHisTyrProLysPheGlyValSerGluAspCyLeu 128
|||
DB 370 AACGAGAAGGAACAATTCTTACAGTT-----TGTGAAGCTGCTTC 414
QY 129 TyrLeuAsnLeuTyrLysProAlaHisAlaAspThrGlySerLysLeuProValLeuVal 148
|||
DB 415 TACCTGAATATTTACACTCTGCTGACTTGACAAAGAACAGCAGCTACACAGATGATG 474
QY 149 TrpPheProGlyValAlaPheLysThrGlySerAlaSerLysAspGlySerAlaLeu 168
|||
DB 475 TGGATCCATGGAGTGGAGCTGCTGCTGGGCGGAGCATCCACTATATGATGATCGGCCCTC 534
QY 169 AlaAlaTyrGluAspValLeuValValValGlnTyrArgLeuGlyLysPheGlyPhe 188
|||
DB 535 TCTGCCCATGAAATATGCTGCTGCTGATGACCATTCAGTATCGCTTGGCATCGGGGATTC 594
QY 189 PheThrThrTrpAspGlnHisAlaPheProGlyAsnTrpAlaPheLysAspGlnValAlaAla 208
|||
DB 595 TTCAGCACTGGGATGAAACAGCTCGGGAAACTGGGCTCACTTGACCAAGTGGCTGCA 654
QY 209 LeuSerTrpValGlnLysAsnIleGluPheGlyGlyAspProSerSerValThrIle 228
|||
DB 655 CTACGCTGGGTCAAGAACAACTTGGGAGCACTTGGGGGCAACCAAGCTCGGTGACATC 714
QY 229 PheGlyGluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLys 248
|||
DB 715 TTTGAGAGATCGACGAGAGTTCAGTCTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCC 774
QY 249 GlyLeuPheHisLysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAla 268
|||
DB 775 AACCTTCTCAAGGCGCACTTCTAGATGCTGATGCTGCTCACTGCTGCTGATTAACA 834


```

Oy HisaPryrglulyserserleuaspheuglnvalalanhlsphesCyeglyAsnAsnla 288
Db ACAGATGTAAGCCCATTCCTGCTG-----GTTGCACTCTTTCGGGTAAACT 888
Oy SeraspSerGluAlaLeuleuArgCysleuagthrlvProserLygluleuLeuthr 308
Db ACCACATCAGCTGTATGATTCATTCCTGCCAGAAAGACAGAGATGMACTACTGAG 948
Oy LeuSerGlnlyrThlyserPhe----- 316
Db ACCTCACTAAATTGAATCTTTTAACTGCACTTACCTGAATCCAAAGAGACTAT 1008
Oy -----ThrArgValValAspGlyAlaPhePheProAsnGluProleuaspheuleu 333
Db CCCTTCCTCCCTACCTGATATGATGAGTATGCGCAAGAGCACAGAAAGATCTCG 1068
Oy SerGlnlyrAlaPheLyAlaIleProserIleIleGlyValAsnAsnIleGlyCysgl 353
Db GCTGAGAAAGATTTCAGACCTGCTCCCTACATAGTGGCATCAACAGCAGAGTTTGGC 1128
Oy PheLeuLeuPrometLyeglulAlaProGluIleLeuSerGlySer---AsnlySerleu 372
Db TGGATCATTCCTCAACCTTATGCGCTATCCACTGCTGAAGCAAACTGGACCAAGACA 1188
Oy AlaLeuHisLeuIleGlnAsn-----IleuHisIleProProGlnlyrLeuHis 389
Db GCCATTCCTCTCTTGTGAAGTCTACCCACCACTTAAATCTCGAAGATGATGATCA 1248
Oy LeuValAlaAsnGlnlyrPheHisAspLyHisSerLeuThrlGluIleArgAspSerleu 409
Db GTGGTCGCTGAGAACTATTAGAGAGACAGATGACCTCACCAAAAGAACCTGTTC 1308
Oy LeuAspLeuLeuGlyAspValAlaPhePheValValProAlaLeuIleThrlAlaArgTrHis 429
Db CAGCATGTGATGGCTGATGTGATTTGGTCCCATGATGATGTGCTCGAAGTCC 1368
Oy ArgAspAlaGlyAlaProVallyrPheTyrgluPheAsnHisArgProGlnCysPheGlu 449
Db AGAGATGCTGAGAGCTCCACCTATATGATGATGATGATGATGATGATGATGATG 1428
Oy AspThrlvAspProAlaPheValIleAspHisIleAspGlyValAlaArgPheValIleGly 469
Db GCCATGAGACCCAGCAGATATAGAGACCATGTGATGATGATGATGATGATGATG 1488
Oy GlyAlaPheLeuLyGlyAspIleValIleMetPheGlnGlyAlaThrlGluGlnlyrleu 489
Db TCTCCATTTTAAA-----GATGGTCTCAGAAAGAGAGACCAAC 1530
Oy LeuSerArgLyMetMetLyserTyrrAlaThrlPheAlaArgThrlGlyAsnProAsnGly 509
Db CTCACGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1590
Oy AsnAspLeuSerLeuTyrrProAlaTyrrAsnLeuThrlGlnGlnlyrleuAspLeu 529
Db GGAGGCTGCCCCACCTGCGCAATATGACAGAAAGAGGATCTGAAGATGATGATG 1650
Oy AsnMetSerLeuGlyGlnArgLeuLyglulProAlaGlyAlaPhePheThrlSerThrlle 549
Db TCAATTCAGGCAAGCCAGAGCTGAGAGCAAGAAAGATGATGATGATGATGATGATG 1704
Oy ProleuIleLeuSerAlaSerAspMetLeuHisSerPro 562
Db -----CTCAGGCGCCAGAGATCAGCCCGAGAGGCCA 1734

```

```

RESULT 12
CF111083 1918 bp mRNA linear EST 23-Jul-2003
LOCUS Shultzemia04334 Rat lung airway and parenchyma cDNA libraries
DEFINITION Rattus norvegicus cDNA clone Contig345 5', mRNA sequence.
ACCESSION CF111083
VERSION CF111083.1 GI:33167595
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)

```

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE

1 (bases 1 to 1918)
Shultz,M.A., Zhang,L., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,
Padua,A.M., Gureke,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,
Plopper,C.G. and Buckpitt,A.R.

TITLE

Gene expression analysis in response to lung toxicants: I.
Sequencing and microarray development

JOURNAL

Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
Contact: Shultz MA
Dept. of Molecular Biosciences, School of Veterinary Medicine
University of California, Davis
1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
Tel: 530 752 0793
Fax: 530 752 4698

COMMENT

Email: mshultz@ucdavis.edu
Average Phred score is 20 or better. All poor quality data (Phred < 20) and vector/linker sequence has been removed.
High quality sequence stop: 1918.

FEATURES

source

1..1918

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="Contig345"

/sex="male"

/tissue_type="airway or parenchyma"

/dev_stage="adult"

/clone_lib="Rat lung airway and parenchyma cDNA libraries"

/note="Organ: lung; Vector: pGSM-112f(-); Site 1: Eco RI;
Site 2: Not I; mRNA was isolated from microdissected rat
lung airways and parenchyma tissues."

ORIGIN

Alignment Scores:

Pred. No.:	1,286-106	Length:	1918
Score:	1097.50	Matches:	239
Percent Similarity:	55.82%	Conservative:	87
Best Local Similarity:	40.92%	Mismatches:	207
Query Match:	35.64%	Indels:	52
DB:	7	Gaps:	9

US-10-023-515-2 (1-581) x CF111083 (1-1918)

```

Oy 5 LeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleGlnProleuGlnlyHis 24
Db 27 CTACAAATGGCTCAACCTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 83
Oy 25 ArgGlnTrpGlyLyserThrlGlyProserAlaGlnGlyProGlnAsnThrlArgleuGly 44
Db 84 -----GGTAT-CCATCTCACCACTGTGTGTGACACTGTTAAAGGC 124
Oy 45 TrpIleGlnGlyLyserGlnValThrlValLeuGlySerProValAsnValPheLeu 64
Db 125 AAAGTCTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 184
Oy 65 GlyValProPheAlaAlaProProleuGlySerLeuArgPheThrlAsnProGlnProAla 84
Db 185 GGAATCCCTTCGCCAATCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 244
Oy 85 SerProTrpAspAsnLeuArgGluAlaThrlSerTyrrProAsnLeuCyLeuGlnAsn--- 103
Db 245 GAGTCTTGAACTTTGTGAGAAATATCTACTCTTACCCACCTGATGTCTTCAAGATCT 304
Oy 104 -----SerGluTrpLeuLeuAspGlnHisMetLeuLyValHis 117
Db 305 GTTGAGGGCAGGTTCTCTCAGAGCTTTTCAACCAAGAGAAAGAAATTCCTTTACAG 364
Oy 118 TyrrProLyPheGlyValSerGluAspCysLeuTyrrleuAsnIleTyrrAlaProAlaHis 137

```

```

Db      365 TTT-----TCGAAGACTGCGCTACTGATGATTACTCCCGCTGAC 409
Qy      138 Alaasphrctglyserlyseuprovalleuvaltrp pheprogl ygl yalaphelythr 157
Db      410 TTGACAAAGAACCGCGTTACCAAGTATGATGATCCATCCAGAGGTGACGTATG 469
Qy      158 Glyseralserllephepserglyseralaleualalatrpglyuapvalleuval 177
Db      470 GGTGATCATCCACCTATGATGACAGGCGCTCTGCGCCATATAAATGGGTGGTG 529
Qy      178 Valvalglntfaryglleuglyllepheglyphepethrthrtrpaspghnlsalapro 197
Db      530 ACCATTGAGATGATGCGCTTGGCATCTGGGGATCTTCAGACACAGGGATGAAACACAGCCAG 589
Qy      198 GlyAsntrpAla pheLyAspGlnValAlaAlaLeuSerTrpValGlnlyAsnIleGlu 217
Db      590 GGCACCTGGGGTATTGTCACAGCTGGCTGACCTGAGGATGAGGATTCACAGACATATATGCGC 649
Qy      218 PhepheglylyAspProserSerValtrpIlePheglylyuSerAlaGlyAlaIleSer 237
Db      650 AACTTGGGGGCACACCGAGCTCTGACCATCTTGGAGATCTGCAGAGGATTCAGT 709
Qy      238 ValSerSerleuIleLeuSerProMetAlalyGlyLeuPhehlyValIleMetGlu 257
Db      710 GTCTCTGTTCTTGTATTCTCTCTGCGCAAGACCTCTCCACAGGGCCATTCTGAG 769
Qy      258 SerGlyValAlaIleleProtyrleuglyuAlaHisAspTyrcyluSerleuSerleu 277
Db      770 AGTGGGTGTCCTCTCAATCTGCTCTGATTACACAGATGACAGAGCCATTGCTAATGCG 829
Qy      278 GlnValValAlaHisPheCyGlyAsnAsnAlaSerAspSerGlyuAlaLeuLeuArgCys 297
Db      830 -----ATTGCTACTCTTCTGGGTGTAAACACACACATCAGCTTATGATGCTTCAATGCG 883
Qy      298 LeuArgThrLyPProSerLyGlyLeuLeuThrLeuSerGlnlyThrLySerPhe--- 316
Db      884 CTGCCCGAGAACAGAGGATTAATCTCTGAGACTTCAATTAATTAATCTTTTAA 943
Qy      317 -----ThArgValAlaAspGly 322
Db      944 CTGACACTATTGGAACCAAAAGAGCTATCCCTTCCTACCTACCTGATTTGATGGA 1003
Qy      323 AlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnlyAlaPheLyAlaIlePro 342
Db      1004 GTGCTGCTGTAAAGACACAGAGAGATCTGCTGAGAGAGTTTCAACAGCGTCCC 1063
Qy      343 SerIleIleGlyValAsnAsnHisGlyCysGlyPheLeuLeuProMetLyGluAlaPro 362
Db      1064 TACATAGTGGGCATCAACAGCAGAGATTGGCTGATCATTCCAATGCTTATGGGCTAT 1123
Qy      363 GlnIleLeuSerGlySer--AsnLySerLeuAlaLeuHisLeuIleGlnAsn----- 379
Db      1124 CCACTCTCCGAGGCAAACTGACACAGAAACAGCAATTCCTCTTGAGAGTCCCTAC 1183
Qy      380 ---IleuHisIleProProGlnTyrlleuHisleuValAlaAsnGlnTyrlPheHisAsp 398
Db      1184 CCAACACTGAAATCTCTGAGAAATATGATTCATAGGCTCTTGAGAGTATTTAGAGAGG 1243
Qy      399 LyHisIleSerleuThrGlnIleArgAspSerleuLeuAspLeuGlyAspValPhePhe 418
Db      1244 ACAAGATGACCTGCGAAGAGAAAGACCTGTTCAAGACTTGGTGCAGATGTATTTT 1303
Qy      419 ValValProAlaLeuIleThrAlaArgTyrlHisArgAspAlaGlyAlaProValTyrlPhe 438
Db      1304 GGTGTCCTCATCGATGATGCTCTGAGAGTCAACAGATGCTGAGAGCCCACTTCATG 1363
Qy      439 TyrlGlnPheArgHisArgProGlnCysPheGlnAspThrLyAspProAlaPheValLyAspAla 458
Db      1364 TATGAAATTTAGATATCCCAAGCTTGTATCAGCCATGAGGCCCAAGACAGATGCGGA 1423
Qy      459 AspHisAlaAspGlnValArgPheValPheGlylyAlaPheLeuLyGlyAspIleVal 478
Db      1424 GACCATGATGATGAAATCTCTCAGTATTTGATCTCATTTTAA----- 1471

```

```

Qy      479 MetPheGlnGlyAlaThrGlnGlnGlyLeuLeuSerArgLyMetMetLySerTrp 498
Db      1472 -----GATGGGCGCTCGAAGAGAGACCAATCTCAGCAAAATGATATAATACG 1525
Qy      499 AlaThrPheAlaArgThrGlyAsnProAsnGlnlyAsnAspLeuSerleuTrpProAlaTy 518
Db      1526 GCAACTTGTCTCGGAATGGGAACCTTAATGGGGAGGCGTCCCAATTGGCCAGAAATAT 1585
Qy      519 AsnLeuThrGlnGlnTyrlleuGlnLeuAspLeuAsnMetSerleuGlyAlaArgLeu 538
Db      1586 GACCAAGAGAGAGGATCTTAAGATTTGCTCTCACTCAGGACGCCAGAGGCTGAG 1645
Qy      539 GlnProArgValAlaPhePheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMet 558
Db      1646 GACAAAGAAATGGCTTTTGTGCTGAG-----CTCAAGGCCCAAGAGAGCA 1690
Qy      559 LeuHisSerPro 562
Db      1691 GCAAGAGAACCA 1702

RESULT 13
CR618303
LOCUS full-length cDNA clone CS0DE014YC12 of Placenta of Homo sapiens
DEFINITION (human).
ACCESSION CR618303
VERSION CR618303.1 GI:50499110
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1826)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1826)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen
FEATURES
source
1..1826
location/Qualifiers
1..1826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE014YC12"
/rna_processing="cDNA"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 1,246-103 Length: 1826
Score: 1069.50 Matches: 224
Percent Similarity: 57.60% Conservative: 79
Beet Local Similarity: 42.59% Mismatches: 142
Query Match: 34.74% Indels: 82
Gaps: 6
US-10-023-515-2 (1-581) x CR618303 (1-1826)
Qy      30 ThrGlyProSerAlaGlnGlyProGlnArgAsnThrArgLeuGlyTyrlIleGlnGly 49

```

```

Db      84  ACTGGCCCGAAGTTGCTCAGCCTGAGATGACACACCCCTGGCTGTCGAGGCGCG 143
Qy      50  GlnValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 69
Db      144  CAGGTGGGGGTAAAGGACACAGCCGCTTGTAAGTCTTCTGGGCACTTCATTTGGCC 203
Qy      70  AlaProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsn 89
Db      204  CAGCCGCCACTGGGCGCTGACCGGTTCTAGGCCACACACCGACAGCCAGCCCTGGAGGCT 263
Qy      90  LeuArgGlyAlaThrSerTyrrProAsnLeuCyLeuGlnAsnSerGluTrpLeuLeu 109
Db      264  GTGCCGGATGCGCACCTGCGCCCAATGTCTTACAGACGATGAGACATGACACAGC 323
Qy      110  AspGlnHisMetLeuLysValHisTyrrProLysPheGlyValSerGluAspGlyLeuTyrr 129
Db      334  AGCAAGATTTGCTCTCAACGGAAACAGACAGATCTTCTCCGTTTCAGAGAGACCTGCTGTC 383
Qy      130  LeuAsnIleTyrrAlaProAlaHisAlaAspThrGlySerTyrrLeuProValIleValTrp 149
Db      384  CTCACGCTCTATAGCCCAAGCTAGAGTCCCGGAGGGTCCGTAAGCCGCTCATGGTATGG 443
Qy      150  PheProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAla 169
Db      444  GTCCATGAGAGCGCTCTGATACTGGCGCTGCCACCTCTACAGATGATCAGCTCGGCT 503
Qy      170  AlaTyrrGluAspValLeuValValValGlnTyrrArgLeuGlyIlePheGlyPhePhe 189
Db      504  GCTTATGGGAGATGTGTCTGTGTACAGTACAGTACCGCTTGGGCTCTTGGCTTCTTTC 563
Qy      190  ThrThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeu 209
Db      564  AGCATCTGAGATGAGATGACATGACCTGGCAACAGGGCTTCTAGATGTGTAGCTCTTTC 623
Qy      210  SerTrpValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePhe 229
Db      624  CGCTGGGTGCAAGAAACATCGCCCTTCGGGGGTGACCTCACTGTCTCATCTGCTT 683
Qy      230  GlyGluSerAlaGlyValAlaSerValSerSerIleLeuSerProMetAlaLysGly 249
Db      684  GGTGATCTGCGCGGAGCATCTCTGCGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTG 743
Qy      250  LeuPheHisLysValIleMetGluSerGlyValAlaIleIlePro---TyrrLeuGlnAla 268
Db      744  CTGTTTCCACAGGCCATTCACACAGATGGAGTACATCACACCCAGGATCATTCAGCTT 803
Qy      269  HisAspTyrrGluLysSerGluAspLeuGlnValAlaHisPheCysGlyValAsnAla 288
Db      804  CACCTTGGCCCTTACGCT-----CAGAAATCGCAACACCTTGGCTGCAAGCTCC 854
Qy      289  SerAspSerGlyAlaLeuLeuArgCyLeuAlaGlyThrLysProSerLysGluLeuThr 308
Db      855  AGCTCCCGGCTGAGATGTGTGCACTGCTTCCAGCAAGAAAGAGAGAACGCTGCTC--- 911
Qy      309  LeuSerGlnLysThrLysSer-----PheThrArgValValAspGlyValAlaPhePro 326
Db      912  CTTACCAAGAAAGCTGAAATATCTATCTTCTCTCACCGTTGATGAGCTGTCTTCCCC 971
Qy      337  AsnGluProLeuAspLeuLeuSerGlnLysValAlaPheLysAlaIleProSerIleIleGly 346
Db      972  AAAAGCCCAAGAGAACTCTGAAAGAGAAAGCCCTTCACTGTGCTCTCTCATGGT 1031
Qy      347  ValAsnAsnHisGlyCysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSer 366
Db      1032  GTCAACAACACATGAGTTCAGCTGCTCATGCCAG----- 1066
Qy      367  GlySerLeuLysSerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGln 386
Db      1066  ----- 1066
Qy      387  TyrrLeuHisLysLeuValAlaAsnGluTyrrPheHisAspLysHisSerLeuThrGluIleArg 406
Db      1066  ----- 1066

```

```

Qy      407  AspSerLeuLeuAspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAla 426
Db      1066  ----- 1066
Qy      427  ArgTyrrHisArgAspAlaGlyAlaProValTyrrPheTyrrGluPheArgHisArgProGln 446
Db      1067  -----GATCTTGAGAGCCCTGTCTTCTTCTATGATTCAGATCAGACCACT 1114
Qy      447  CysPheGluAspThrLysPheProAlaPheValLysAlaAsnHisAlaAspGluValArgPhe 466
Db      1115  TCTTTTGGCAAGATCAAACTGCTGCGTGAAGGCTGATCAAGGGCCAGGGTCTCTTT 1174
Qy      467  ValPheGlyGlyAlaPheLeu-----LysGlyAspIleValMetPheGluGlyAlaThr 484
Db      1175  GTGTTGAGAGTCCCTTCTCTATGACACAGAGCTCCCGCTGCTTTCAGAGGCCACA 1234
Qy      485  GluGluGluLysLeuLeuSerArgLysMetLysTyrrTrpAlaThrPheAlaArgThr 504
Db      1235  GAGAGAGAGAGACACTAAGCTTCAACATGATGGCCAGATGGACCACTTTCGCGACA 1294
Qy      505  GlyAsnProAsnGlyValAsnAspLeuSerLeuTrpProAlaTyrrAsnLeuThrGluIleTyrr 524
Db      1235  GGGGACCCCAATAGCAAGGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1354
Qy      525  LeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLysGluProArgValAspPhe 544
Db      1355  CTGGAGATCAACCCAGCTGCCACAGGCGGACGACAGAAATTCAGGAGAGCCCTGATGCATTC 1414
Qy      545  TrpThrSerThrIlePro 550
Db      1415  TCGTCAGAGAGCGCTCCCC 1432

```

```

RESULT 14
CRE61368      1859 bp      mRNA      linear      HTC 18-AUG-2004
LOCUS      CRE61368
DEFINITION      Tetradonon nigroviridis full-length cDNA.
ACCESSION      CRE61368
VERSION      CR61368.1 GI:51137813
KEYWORDS      HTC; cDNA; full-length; Tetradonon nigroviridis.
SOURCE      Tetradonon nigroviridis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoleleostei;
      Tetradontidae; Tetradontidae; Tetradonon.
REFERENCE      1 (bases 1 to 1859)
AUTHORS      Genoscope.
TITLES      Direct Submision
JOURNAL      Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
      ; 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
      (E-mail : segrefigenoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT      The sequences are based on single pass reads.
      More information available at
      http://www.genoscope.cns.fr/tetradonon.
FEATURES
      source
      1..1859
      /organism="Tetradonon nigroviridis"
      /mol_type="mRNA"
      /db_xref="taxon:99883"
      /cissue_type="Liver"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      4,67e-102      Length:      1859
Score:          1055.00      Matches:      228
Percent Similarity: 59.70%      Conservative: 89
Beet Local Similarity: 42.94%      Mismatches: 184
Query Match:    34.26%      Indels:      30
DB:             3      Gaps:      10

```

US-10-023-515-2 (1-581) x CRE61368 (1-1859)

```

QY 37 ProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLeuValThrValLeuGlySer 56
Db 113 CCGTAATTCATACCAAACTGGACAGCTGAGAGTGATATCAAGACGGTAAGGGGAAA 172
QY 57 ProValProValAsnValPheLeuGlyValProPheValAlaProPheLeuGly---Ser 75
Db 173 GAGCGTGAAGTCCATGATACCTGGGTGTGCTCTTCTGTAAGCACTCGGGTCCGTCT 232
QY 76 LeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArgGlyAlaThrSer 95
Db 233 CTGAGCGTGGCTCCACCCAGCCCTGTAGAGCGCTTGGAGAGAGATGCTACCAAG 292
QY 96 TyrProAsnLeuCyLeuGlnAsnSerGlu-----TrpLeuLeuAspGlnHisMet 113
Db 293 CAACCGGTAGTGTGTGTGTCAGATTTAGAGTTGCAACACAAATATATTAAGAACTTAAT 352
QY 114 LeuLeuValHisTyrProIlyAspPheGlyValSerGluAspCyLeuTyrLeuAsnIleTyr 133
Db 353 TTAGAAAGTCAACCTTCA-----GACATTTCAAGAAAGCTGCTTAACCTCAACATTTAC 406
QY 134 AlaProAlaHisAlaAspThrArgSerIlySerIlyLeuProValLeuValThrPheProGlyGly 153
Db 407 ACTCGCTGCCAAGACAGCTGAGAAATACCAAGCTCCAGTTATGTGTGTGATCCACGGTGA 466
QY 154 AlaPheIlyThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIatyrGluAsp 173
Db 467 GGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 526
QY 174 ValLeuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThrTrpAsp 193
Db 527 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 586
QY 194 GlnHisAlaProGlyAsnTrpAlaPheIlyAspGlnValAlaAlaLeuSerTrpValGln 213
Db 587 AAGCAGATCCAGCAAAACATGTGCTTGTGATCAAGCTCAAGCTCTGAAAGGGTCCAG 646
QY 214 LysAsnIleGlnPhePheGlyIlyAspProSerSerValThrIlePheGlyIlySerAla 233
Db 647 GAGCAATATTCACAACTTGTGAGAGAACCCAGATTTAGTACATATTTGGCGAGTCTCT 706
QY 234 GlyAlaIleSerValSerSerIleLeuSerProMetAlaIlyGlyLeuPheHisIly 253
Db 707 GGTGAGTGAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 766
QY 254 AlaIleMetGlySerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyrGlyIly 273
Db 767 GCAATTTGCTGAAGAGTGCACCTGCTCAATG---CATATATTTGCTCAGAT---GACCC 820
QY 274 SerGluAspLeuGlnValAlaHisPheCyAspGlyAsnAlaSerAspSerGluAla 293
Db 821 CAGCCAGTGTACAGCTTGTGCAAAATGATATCTGTTGTACACCTGAAAGCAGAAAG 880
QY 294 LeuLeuArgCyLeuArgThrIlySerProSerIlyGlyLeuLeuThrLeuSer---GlnIly 312
Db 881 ACTGCTGATTCATCAAGAACCTGAGAGTGAAGCAATTTTACCTTTTACCTAAGAAAAA 940
QY 313 ThrIlySerPheThrArgValIlyAspGlyAlaPhePheProAsnGluProLeuAspLeu 332
Db 941 GAGTTAAGATACCTTAATATCAGATGATATTTCTCAAAAATGACAGAGAAAGTTG 1000
QY 333 LeuSerGlnIlyValAlaPheIlyAlaIleProSerIleIleGlyValAlaAsnHisGlnIly 352
Db 1001 TTCTGTAGAGATCAAGATTTCTCACTGTCTTTCATGACTGAGTCAATTAATTAAGAGT 1060
QY 353 GlyPheLeuLeuProMetIlyGlyAlaPro-----Gln 363
Db 1061 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1120
QY 364 IleLeuSerGlySerAlaIlySerLeuAlaLeuHisIleLeuIleGlnAlaIleLeuHisIle 383
Db 1121 CAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1153
QY 384 ProPro-----GlnTyrLeuHisIleLeuValAlaAsnGlnTyrThrPheHisAspIlyHis 400

```

```

Db 1154 GATCCCAAGATGACAGCTTAGAAAGAGGATTTTAATGATATATACGAACCTGGGAG 1213
QY 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
Db 1214 GATCTGTGAAGAAACAGAGACGGGTACACTGAGGGTCTTGGAGACATATATTTGTCAATT 1273
QY 421 ProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGlu 440
Db 1274 CCGCCATTAAGAGCTGTATTTGCTTACAGATGAGAGTGTCTCTGTCTACTGTATAGAG 1333
QY 441 PheArgHisArgProGlnCyPheGluAspThrIlySerProAlaPheValIlyAlaAspHis 460
Db 1334 TTTCAGCATTTCCCCAGTTTCTTAAGGATTAAGATCAACCTTGTATGAGAGTACCAT 1393
QY 461 AlaAspGluValAcArgPheValPheGlyGlyAlaPheLeuIlyGlyAspIleValMetPhe 480
Db 1394 GAGATGAAATATATTTTGTATTTGGGTTTGTTCACAGTTTCTCACTCAATCATTAAGC 1453
QY 481 GluGlyAlaThrGluGluGluIlyLeuLeuSerArgIlySmetIlySlyTyrPalaThr 500
Db 1454 AGTGGTGCACATGAAGAGAGAGACAGATTGACCCAAATCATGATGAGCTACTGGGAGAC 1513
QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpAlaIlyTyrAsnLeu 520
Db 1514 TTGTGTGCACAGGGGTCTCTTATGTGACAGGGTCTTTCACCTGGCCAAAGTATGAGAA 1573
QY 521 ThrGlnTyrTyrLeuGlnLeuAspLeu-----AsnMetSerLeuGlyGlnArgLeuGlu 539
Db 1574 ACGAGAGAGTACTCGGGAATTGATGCAAGAGAGCATTTTATCATCGGTTTGAAAAA 1633
QY 540 ProArgValAspPheTrpThrSerThrIlePro 550
Db 1634 GACAAATATCTCTCTTACCCACAGACCCCA 1666

RESULT 15
CR634885 1817 bp mRNA linear HTC 18-AUG-2004
LOCUS
DEFINITION Tetraodon nigroviridis full-length cDNA.
ACCESSION CR634885
VERSION GI:51131330
KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 1817)
REFERENCE
AUTHORS Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
TITLE Direct Submission
JOURNAL
COMMENT
1 2 rue Gaston Crémieux, CP 5706 - 91057 Evry cedex - FRANCE
(E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source
1..1817
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/issue_type="Liver"

ORIGIN
Alignment Scores:
Pred. No.: 2,566-101 Length: 1817
Score: 1048.00 Matches: 227
Percent Similarity: 59.51% Conservatve: 89
Best Local Similarity: 42.75% Mismatch: 185
Query Match: 34.04% Indels: 30
DB: 3 Gaps: 10

```

US-10-023-515-2 (1-581) x CR634885 (1-1817)

```

Qy 37 ProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyValIleThrValLeuGlySer 56
Db 93 CCGAAATCCATCAACCACTGGCAGCTGAGAGGGATATACAGACCTGAAGGGGAAA 152
Qy 57 ProValProValAsnValPheLeuGlyValProPheAlaAlaProProLeuGlySer 75
Db 153 GACGCTGAGTCAATGATACCTGGGTGGCCCTTTGCTAACACCACTCTGGTCTCTGCT 212
Qy 76 LeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeuArgGlyAlaThrSer 95
Db 213 CTGAGGCTGGCTCCACCCAGCCTGTAGAGGCTTGGAAGAGAGATGATCTACCAAG 272
Qy 96 TyrProAsnLeuValLeuGlnAsnSerGlu-----TrpLeuLeuLeuAspGlnIleMet 113
Db 273 CAACCGCTGATGTGTGTTGAGATTGAGGTTGCAACCAAAATTAATTAACACTTAAT 332
Qy 114 LeuValValIleValIleValIleValIleValIleValIleValIleValIleValIle 133
Db 333 TTGAAAGTCAACCTTCA-----GACATTTCAGAAAGCTGCTTACCTCAACATTTC 386
Qy 134 AlaProAlaAlaIleAlaAspThrGlySerIleValLeuProValIleValIleValIle 153
Db 387 ACTCTGCAACAGACAGCTGAGAAATACCAAGCTCCAGATTATGCTCGATCCAGGTGA 446
Qy 154 AlaPheValIleValIleValIleValIleValIleValIleValIleValIleValIle 173
Db 447 GGGTTTGTATTAGGTTGGCTTCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 506
Qy 174 ValLeuValIleValIleValIleValIleValIleValIleValIleValIleValIle 193
Db 507 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 566
Qy 194 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 213
Db 567 AAGCATAGCCAGAGAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 626
Qy 214 LysAsnIleGlnPhePheGlyIleValAspProSerValIleIlePheGlyIleLeuSer 233
Db 627 GAGCATATTCAACACTTGGAGAGAGACCAAGTTTAGTACCAATTTGGCCAGCTGCT 686
Qy 234 GlnAlaIleSerValSerSerIleLeuSerProMetAlaIleGlyLeuPheIleVal 253
Db 687 GGTGAGTGAAGTGTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
Qy 254 AlaIleMetGlnSerGlyValAlaIleIleProTyrIleGlnAlaIleAspTyrGly 273
Db 747 GCAATTGCTGAAAGTGGCAGCTGCTCAATG---CATATATTGCTCAAGAT---GACCC 800
Qy 274 SerGlnAspLeuGlnValIleAlaIlePheCysGlyAsnAsnAlaSerAspSerGly 293
Db 801 CAGCCAGTGTAAACATTTGTGCAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
Qy 294 LeuLeuArgCysLeuArgThrIleValAspProSerIleValLeuLeuThrLeuSer---Gln 312
Db 861 ACTGCTGATTCATAGAACTGGAAGCTGGAACCACTTTTACCTTAAGGAAATA 920
Qy 313 ThrIleSerPheThrArgValIleValIleValIleValIleValIleValIleValIle 332
Db 921 GAGTTAAGTATCTTATATACCGATGATTTCTCAAAATAATCAGCAGAAAGTTG 980
Qy 333 LeuSerGlnValIleValIleValIleValIleValIleValIleValIleValIleVal 352
Db 981 TTCCTTGAGCATGAATCTTCACTGCTGCTTTCATGACTGAGTCAATATATATAGAGGT 1040
Qy 353 GlnPheLeuLeuProMetIleValIlePro-----Gln 363
Db 1041 GGTGTGTTACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
Qy 364 IleLeuSerGlySerAsnIleValIleValIleValIleValIleValIleValIleVal 383
Db 1101 CACGCTCTGGT-----TTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133

```

```

Qy 384 ProPro-----GlnTyrIleuIleValIleAlaAsnGluTyrPheIleAspIleVal 400
Db 1134 GATCCCAAGATGACAGCTTATGAAAACGATTTCTTAATGATATATACAGAACTGGGAG 1193
Qy 401 SerIleuThrGlnIleArgAspSerIleuAspLeuAspLeuGlyAspValPhePheVal 420
Db 1194 GATCGTGGAAAACAGAGACGGGTACCTGAGAGGTGCTTGGAGCATGATATTTGTCATT 1253
Qy 421 ProAlaLeuIleThrAlaArgTyrIleArgAspAlaGlyAlaProValTyrPheTyrGlu 440
Db 1254 CCTGCAATTAAGCTGCTATGCTTCAAGAGATGAGAGGTGCTGCTGCTGCTGCTGCTGCT 1313
Qy 441 PheArgIleArgProGlnCysPheGlnAspThrIleValPheValIleValIleValIle 460
Db 1314 TTGCAGCATTTCCCAAGTTTCTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 1373
Qy 461 AlaAspGlnValIleArgPheValPheGlyIleValIleValIleValIleValIleVal 480
Db 1374 GGAATGAAATATATTTTGTATTTGGCTTTGTTTCAAGTTTCTCAATCAATCAATCAAT 1433
Qy 481 GlnGlyAlaThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 500
Db 1434 AGTGCCTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1493
Qy 501 PheAlaArgThrGlnValAsnProAsnGlyAsnAspLeuSerIleuProAlaTyrAsnLeu 520
Db 1494 TTGCTCCACAGGGTCTCTTAATGACAGGGCTTTGTCACAGTTCAGCAAGTATGAGAGA 1553
Qy 521 ThrGlnGlnTyrIleuGlnLeuAspLeu---AsnMetSerLeuGlyIleArgLeuVal 539
Db 1554 ACGGAGAGTACCTGGAATGATGACAAAGAGCAAGATGTTATCAAGGTTTGAATAA 1613
Qy 540 ProArgValAspPheTrpThrSerThrIlePro 550
Db 1614 GACAAATATCTCTCTTCAACCCAGACCCCCCA 1646

```

Search completed: June 15, 2005, 13:29:40
 Job time : 5297 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 15, 2005, 09:44:41 ; Search time 283 Seconds

(without alignments)
3359.282 Million cell updates/sec

Title: US-10-023-515-2

Perfect score: 3079

Sequence: 1 MPQGLTSSASQWCFLLIQP.....PLSLTFLSLQPPFFFCAP 581

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 81813835 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh
-O=/cgn2_1/USPRO.spool/6664091/runat_14062005_133641_10487/app_query.fasta.1.775
-DB=Issued Patents NA -QFMT=fastcap -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPCL=0
-DBEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=psco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091_0CGN_1_1_69_0runat_14062005_133641_10487 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3079	100.0	1746	4	US-10-023-515-3	Sequence 3, Appli
2	3079	100.0	2158	4	US-10-023-515-1	Sequence 1, Appli
3	1230	39.9	2117	4	US-09-949-016-3799.	Sequence 3799, Ap
4	1230	39.9	2169	4	US-09-949-016-555	Sequence 555, App
5	1230	39.9	2191	4	US-09-595-682B-27	Sequence 27, Appl
6	1140.5	37.0	1717	4	US-09-595-682B-20	Sequence 20, Appl
7	1138.5	37.0	1701	4	US-09-264-737-3	Sequence 3, Appli
8	1129.5	35.7	965	4	US-09-799-451-155	Sequence 155, App
9	793	25.8	2802	4	US-09-949-016-2515	Sequence 2515, Ap
10	789.5	25.6	4960	4	US-09-949-016-1017	Sequence 1017, Ap
11	782	25.4	2862	4	US-09-949-016-2516	Sequence 2516, Ap
12	775	25.2	4233	4	US-09-491-356C-5	Sequence 5, Appli

13	764.5	24.8	1845	1	US-07-732-962A-1	Sequence 1, Appli
14	764.5	24.8	1845	5	PCT-US92-06106-1	Sequence 1, Appli
15	764.5	24.8	2158	4	US-09-949-016-1192	Sequence 1192, Ap
16	764.5	24.8	2158	4	US-09-949-016-1193	Sequence 1193, Ap
17	764.5	24.8	2256	2	US-08-318-826A-5	Sequence 5, Appli
18	764.5	24.8	2256	2	US-08-370-156-1	Sequence 1, Appli
19	764.5	24.8	2256	2	US-08-814-095-1	Sequence 1, Appli
20	759.5	24.7	3096	2	US-08-318-826A-6	Sequence 6, Appli
21	759.5	24.7	3096	2	US-08-370-156-3	Sequence 3, Appli
22	759.5	24.7	3096	2	US-08-814-095-3	Sequence 3, Appli
23	759.5	24.7	5767	4	US-09-810-861B-3	Sequence 3, Appli
24	759.5	24.7	14446	4	US-09-810-861B-4	Sequence 4, Appli
25	755.5	24.5	1725	4	US-09-810-861B-5	Sequence 5, Appli
26	755.5	24.5	3016	2	US-08-318-826A-7	Sequence 7, Appli
27	755.5	24.5	3016	2	US-08-370-156-5	Sequence 5, Appli
28	755.5	24.5	3016	3	US-08-814-095-5	Sequence 5, Appli
29	755	24.5	2400	6	5215909-13	Patent No. 5215909
30	755	24.5	2400	6	5215909-13	Patent No. 5215909
31	752	24.4	2381	2	US-08-318-826A-9	Sequence 9, Appli
32	752	24.4	2416	2	US-08-318-826A-8	Sequence 8, Appli
33	752	24.4	2416	3	US-09-334-489-1	Sequence 1, Appli
34	752	24.4	2416	3	US-09-334-489-2	Sequence 2, Appli
35	752	24.4	2446	4	US-09-949-016-5275	Sequence 5275, Ap
36	677	22.0	6667	4	US-09-949-016-17017	Sequence 17017, A
37	673.5	21.9	1764	4	US-09-491-356C-3	Sequence 3, Appli
38	664.5	21.6	4436	4	US-08-445-050-8	Sequence 8, Appli
39	660	21.4	2184	1	US-08-204-691-8	Sequence 8, Appli
40	660	21.4	2184	1	US-08-204-691-8	Sequence 8, Appli
41	660	21.4	2344	4	US-09-347-878-31	Sequence 31, Appli
42	660	21.4	2375	4	US-09-949-016-3976	Sequence 3976, Ap
43	660	21.4	2428	1	US-08-445-050-1	Sequence 1, Appli
44	660	21.4	2428	1	US-08-204-691-1	Sequence 1, Appli
45	660	21.4	2428	4	US-09-355-295B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-023-515-3

Sequence 3, Application US/10023515

Patent No. 6664091

GENERAL INFORMATION:

APPLICANT: Curtis, Rory A. J.

APPLICANT: Silos-Santiago, Immaculada

TITLE OF INVENTION: 53010. A NOVEL HUMAN CARBOXYESTERASE

TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF

FILE REFERENCE: 10448-122001

CURRENT APPLICATION NUMBER: US/10/023,515

CURRENT FILING DATE: 2001-12-18

PRIOR APPLICATION NUMBER: 60/256,369

PRIOR FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: 60/279,508

PRIOR FILING DATE: 2001-03-28

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1746

TYPE: DNA

ORGANISM: Homo sapiens

US-10-023-515-3

Alignment Scores:

Pred. No.: 0 Length: 1746

Score: 3079.00 Matches: 581

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0

US-10-023-515-2 (1-581) x US-10-023-515-3 (1-1746)

QY 1 MetProGInGlyLeuThSerSerzLaSerGInTPCySphePheLeuIleLeuGInPro 20

|||||


```
Db      1  ATGCCACAGAGGACTTACTTATCTGCTTACCAATGATGGCTTTTCTTGATTTCTCCAGCCC 60
Qy      21  LeuLeuGlnYHSAAGGINTTPGILYVrThrGlyProSerAlaGluGlyProGlnArgAsn 40
Db      61  CTGTGGGACACACAGATGGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGAAC 120
Qy      41  ThrArgLeuGlyTPRILeGlnGlyLeuGlnValrThrValLeuGlySerProValProVal 60
Db      121  ACCAGGCTGGGATGATTCAGGGCCAGACCAAGTCACTGTGCTGGAGAGCCCTGTGGCTGTG 180
Qy      61  AsnValrThrLeuGlyValrProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
Db      181  AAGGTGTTCCTCGGAGATCCCCCTTGTGCTGCTCCCCGGCTGGAGTCCCTGGATTTAGAAC 240
Qy      81  ProGlnProAlaSerProTPAAspAsnLeuArgGlnAlaThrSerTyProAsnLeuGly 100
Db      241  CCGCAGGCTGCAATGGCCCTGGAGATCTTGAGAAAGCCACCTTCAACCTTAATTTGTGTC 300
Qy      101  LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnrHISMetLeuValrHISrTyProGly 120
Db      301  CTCACAGACTCAGAGTGGCTGCTTATAGATCAACACATGCTCAAGGTGCAATTCAGGAAA 360
Qy      121  PheGlnValrSerGlnAspGlyLeuTyTrpLeuAsnIleTyTrpAlaProAlaHISAlaAspThr 140
Db      361  TTCGGAGGTCAAGAAAGACTGCTCTACCTGAACATCTATGCGCTGCCACCGCATACA 420
Qy      141  GlySerTyLeuLeuProValLeuValrTrpPheProGlnGlyAlaPheTyThrGlySerAla 160
Db      421  GGCTCCAGACTCCCGCTCTTGGTGGTTCCTCAGAGAGTGGCTTCAAGACTGAGCTCAGCC 480
Qy      161  SerIlePheAspGlySerAlaLeuAlaAlaTyTrpValLeuValrValValGln 180
Db      481  TCCATCTTTGAATGGGTCCGCTGGCTGCTATGAGAGCTGCTGGTGTGTGCTCCAG 540
Qy      181  TyrArgLeuGlyIlePheGlyPhePheThrTrpAspGlnrHISAlaProGlyAsnTrp 200
Db      541  TACCGGCTAGAGATATTTGGTTCTTCAACACATGGAGATCAAGCAATGCTCCGGGAACTCG 600
Qy      201  AlaPheTyAspGlnValrAlaAlaLeuSerTrpValGlnTyAsnIleGlnPhePheGly 220
Db      601  GCCCTCAAGAGACAGATGGCTGCTGCTCGGTCCAGAGAAACATGAGATTTCTTGCT 660
Qy      221  GlyAspProSerSerValrThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
Db      661  GGGAGCCCAAGCTCTGTACCACTTTGAGAGTCCGGGAGCCCATGAAGATGGGGTG 720
Qy      241  LeuIleLeuSerProMetAlaTySerGlyLeuPheHISValrAlaIleMetGlnSerGlyVal 260
Db      721  CTATATCTGTCTCCCATGCGCAAGGCTTATTCACAAAGCCATCAATGAGATGGGGTG 780
Qy      261  AlaIleIleProTyTrpLeuGlnAlaHISAspTyTrpValTySerGlyAspLeuGlnVal 280
Db      781  GCCATCAATCCCTTACCTGGAGGCCCATGATATGAGAAAGTGAAGACCTGACAGTGCTT 840
Qy      281  AlaHISPheGlyValrAsnAlaSerAspSerGlnAlaLeuAlaGlyGlyLeuArgTrp 300
Db      841  GCACATTTCTGTGTACAAATGCTCAGACTCTTAAGGCTCGAGAGTCTCGAGAGACA 900
Qy      301  TyrProSerTyLeuLeuLeuThrLeuSerGlnTyThrTySerPheThrArgValVal 320
Db      901  AAAAACTCCAAAGAGCTGAGCCCTCAGCCAGAAAACAAAGCTTTTCACTCGAGTGCTT 960
Qy      321  AspGlnValrPhePheProAsnGlnProLeuAspLeuLeuSerGlnValrAlaPheTyVal 340
Db      961  GATGGTCTTTCTTCTTAATGAGCTCTGATATCTATGTCTCAGAAAGCATTTAAAGCA 1020
Qy      341  IleProSerIleIleGlyValrAsnAlaHISGlyGlyPheLeuLeuProMetTyGln 360
Db      1021  ATTCCTTCATCATCGAGTCAATTAACAGAGTGGCTTCTGCTGCTTAAGAGAGAG 1080
Qy      361  AlaProGlnIleLeuSerGlySerAsnTySerLeuAlaLeuHISLeuIleGlnAla 380
Db      1081  GCTCTGAGATCTCAGTGGCTCCAAACAGTCCCTTGCCCTCATGTAAACAAACATC 1140
```

```
Qy      381  LeuHISLeProProGlnTrpTyLeuHISLeuValrAlaAsnGlnTyTrpPheHISAspTyHIS 400
Db      1141  CTCACATTCGCCCTCAGATATTTGACCTTGGGTGAATGAATCTTCCACAGACAAAGCAC 1200
Qy      401  SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlnTyAspValrPhePheValrVal 420
Db      1201  TCCCTGACTGAATTCGAGACAGTCTTCTGAGACTTGTGAGATGTCTTCTTGTGGTTC 1260
Qy      421  ProAlaLeuIleThrAlaArgTyHISArgAspAlaGlyValrProValrTyPheTyGln 440
Db      1261  CTTGACATGTACACAGCTCGATATCAACAGATCTGTGTGCACTGTCTACTTCTATAGAG 1320
Qy      441  PheArgHISArgProGlnTyPheGlyAspThrTySerProAlaPheValrValrAlaAspHIS 460
Db      1321  TTTGGGACCGGCTCACTGCTTTTGAACACAGAACCGGCTTTGTCAAGCCGACAC 1380
Qy      461  AlaAspGlnValrArgPheValrPheGlyGlyValrPheLeuTyGlyAspIleValrMetPhe 480
Db      1381  GCTGATGAGTCCGCTTGTGTGCTGCTGAGTGGCTTCTGAAAGGGGACATTTGTTATGCTTC 1440
Qy      481  GlnGlyAlaThrGlnGlnGlnTyLeuLeuLeuSerArgTyMetLeuTyTyTrpAlaThr 500
Db      1441  GAGGAGCCACAGAGAGAGAGATTAAGTCTGAGCCGAAAGATGAATGAAATCTGGGCTACC 1500
Qy      501  PheAlaArgThrGlyAspProAsnGlyAsnAspLeuSerLeuTrpProAlaTyAsnLeu 520
Db      1501  TTTCTCGAACCAGGAAATCTTAATGGGAGACATGCTGTGTGGCCAGCTTATATCTTG 1560
Qy      521  ThrGlnGlnTyLeuGlnIleLeuAspLeuAsnMetSerLeuGlyGlnArgLeuTyGlnPro 540
Db      1561  ACTGAGAGTAACTTCCAGCTGAGCTTGAACATGAGCTTGGAGAGAGACTCAAAAGAACCG 1620
Qy      541  ArgValrAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHIS 560
Db      1621  CGGGTGAATTTTGGACAGACACATCCCTGATCTGTGCTCCACATGCTCCAC 1680
Qy      561  SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePheAspPheCyAla 580
Db      1681  AGTCTCTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy      581  Pro 581
Db      1741  CCT 1743

RESULT 2
US-10-023-515-1
/ Sequence 1, Application US/10023515
/ Patent No. 6664091
/ GENERAL INFORMATION:
/ APPLICANT: Curtis, Roy A. J.
/ TITLE OF INVENTION: Siles-Santiago, Imaculada
/ TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
/ FILE REFERENCE: 10448-122001
/ CURRENT APPLICATION NUMBER: US/10/023,515
/ PRIOR FILING DATE: 2001-12-18
/ PRIOR APPLICATION NUMBER: 60/256,369
/ PRIOR FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: 60/279,508
/ PRIOR FILING DATE: 2001-03-28
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 2158
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (96) ... (1838)
US-10-023-515-1
Alignment Scores:
```

Pred. No.: 0 Length: 2158
 Score: 3079.00 Matches: 581
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-023-515-2 (1-581) x US-10-023-515-1 (1-2158)

QY 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleLeuGlnPro 20
 DB 96 AGCCACAGGAGCTTACCTCATCTGCTTCAACAATGCTGCTTTTCTTCAATTCCTCCAGGCC 155
 QY 21 LeuLeuGlnAlaArgGlnTrpGlyLysThrGlyProSerAlaGlnGlyProGlnArgPhe 40
 DB 156 CTGTTGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGAAC 215
 QY 41 ThrArgLeuGlyLysTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
 DB 216 ACCAGGCTGGATGATTCAGGCAAGCAAGTCACTGCTGGAAAGCCCTGCTGTG 275
 QY 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
 DB 276 AACGGTCTCTGGAGTCCCTTGTCTCCCGCTGGAGTCCCTGGATTACGAC 335
 QY 81 ProGlnProAlaSerProTrpAsnLeuArgGlnAlaThrSerTrpProAsnLeuCys 100
 DB 336 CCGACGCTGATCCGCTGGAGTACCTTCCAGAAAGCCACCTCTTACCTTAATTTGTC 395
 QY 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuLysValHisTrpProLys 120
 DB 396 CTCCAGAACTCAGAGTGGCTGCTTTAGTCAACATGCTCAAGTGTGATACCGAAA 455
 QY 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
 DB 456 TTCCGAGAGTCAAGAACTGCTTACCTTACCATCTAGCCCTGCGCCACCGCAATCA 515
 QY 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
 DB 516 GCTTCACAGCTCCCGCTTGTGAGTTCCTCCAGAGGAGCTTCAAGACTGCTCAGCC 575
 QY 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGluAspValLeuValValGln 180
 DB 576 TCCATCTTTGATGGGTCCGCTGCTGCTATGAGAGCTGCTGTTGTGTCCTCCAG 635
 QY 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpTrpAspGlnHisAlaProGlyAsnTrp 200
 DB 636 TACCGGCTAGGAATTTTGGTTCTTTCACACATGGGATCAGCATGCTCCGGGAACTGG 695
 QY 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGluPheGly 220
 DB 696 GCTTCACAGAGACAGAGTGGCTGCTGCTCCAGAGGAGCTTCAAGAACTCAAGTTCCTCGGT 755
 QY 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
 DB 756 GGGAGCCCAAGCTCTGTACCATCTTTGGCGAGTCCGCGGAGCCATAAGTCTTCTAGT 815
 QY 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisLysValAlaIleMetGlnSerGlyVal 260
 DB 816 CTTATACGTCTCCATGCGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGGG 875
 QY 261 AlaIleIleProTrpLeuGlnAlaHisAspTrpGlyLysSerGlnAspLeuGlnValVal 280
 DB 876 GCCATCATCTCTTACCTGAGGCGCATGATGATGAGAGAGTGAAGCACTGAGAGTGGT 935
 QY 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
 DB 936 GCACATTTCTGTGGTGAACATGCGTCAAGCTGAGGCCCTGCTAGAGGCTTCAAGAGCA 995
 QY 301 LysProSerLysGlnLeuLeuThrLeuSerGlnLysTrpLysSerPheThrArgValVal 320
 DB 996 AAACCTCAAGAGAGCTGCTGAGCCCTCAGACGAGAAAACAAAGTCTTCACTGAGTGGTT 1055

QY 321 AspGlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
 DB 1056 GATGTGGCTTTCTTCTTCCATATGAGCTTACATGATCTATGCTCAGAAAGCATTTAAAGCA 1115
 QY 341 IleProSerIleIleGlyValAsnAsnHisGlyCysGlyPheLeuLeuProMetLysGlu 360
 DB 1116 ATTCCTTCATCATGAGTCAATTAACACAGATGTGGCTTCTCTGCTTGAAGAG 1175
 QY 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLysLeuGlnAsnIle 380
 DB 1176 GCTCTGAGATCTCATAGTGGCTTCAACAGATCTTCCCTTCATCTGATTAACAAACATC 1235
 QY 381 LeuHisIleProProGlnTrpLeuHisLysValAlaAsnGluTrpPheHisAspLysHis 400
 DB 1236 CTGCACATCCGCGCTCAGATTTTGACCTTGGGTATGATATCTTCATCATCAAGAAC 1295
 QY 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
 DB 1296 TCCCTGACAGAAATCCGAGACAGTCTTGGACTTGGAGATGTCTTGTGTGCTC 1355
 QY 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGlu 440
 DB 1356 CTGCACTGATCAAGCTCGATATCAAGAGATGCTGGTGCACCTGTCTACTTCTATGAG 1415
 QY 441 PheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspHis 460
 DB 1416 TTTCGCAACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTGTCAACCGCAAC 1475
 QY 461 AlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
 DB 1476 GGTATGAAATCCGCTTGTGTGCTGCTGCTGCTTCAAGGGGAGCACTGTATGTTTC 1535
 QY 481 GlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpTrpAlaThr 500
 DB 1536 GAAAGAGCCACGAGAGAGAGAGTACTGAGCCGAGAGAGATGAATACTGGGCTAC 1595
 QY 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
 DB 1596 TTGCTCGAACCGGAAATCTTATGGAACAGACTGTCTGTGCGCAGCTTATATCTG 1655
 QY 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
 DB 1656 ACTGAGAGTACCTCAGCTGAGCTTGAACATGAGCTTCGAGCAGAGACTCAAAAGACCG 1715
 QY 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
 DB 1716 CGGGTGAATTTTGGACCAAGCAACATCCCTGATCTGTGCTGCTCCGACATGCTCCAC 1775
 QY 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
 DB 1776 AGTCTCTTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
 QY 581 Pro 581
 DB 1836 CCT 1838

RESULT 3
 US-09-949-016-3799
 / Sequence 3799, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CL001307
 / CURRENT FILING DATE: US/09/949,016
 / PRIOR FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3799
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3799

Alignment Scores:
Pred. No.:      2,14e-136      Length:      2117
Score:          1230.00        Matches:      248
Percent Similarity: 63.84%      Conservative: 91
Best Local Similarity: 46.70%      Mismatches: 168
Query Match:     39.95%      Indels:      24
DB:              Gaps:        8

US-10-023-515-2 (1-581) x US-09-949-016-3799 (1-2117)

QY      31  G|PProSerLagLugLyProGlnAArgAenThrArgLeuGlyTrpIleGlnGlyGln 50
DB      132 GGGCAGGAGACTCAGCCAGTCCCATCCGACACACACAGCGGAGGAGTGGGAGTCTT 191
QY      51  ValThrValLeuGlySerProValProValAenValPheLeuGlyValProPheAlaIa 70
DB      192 GTCCATGTGAAGGCGCCCAATGCCGGGGTCCAAACCTTCCTGGAAATTCATTGGCCAG 251
QY      71  ProProLeuGlySerLeuArgPheThrAenProGlnProAlaSerProTrpAenLeu 90
DB      252 CCACCTCTAGTCCGCTGGATTTGCACTTGGACCCCTCGAATCTTGAGTGTG 311
QY      91  ArgGlnAlaThrSerGlyProAenLeuGlyLeuGlnAen-----SerGln 105
DB      312 AGGAGTGAACCAACCATCCGCGCATGTGTCTACAGAGACCTCACCGCAGTGAAGTCAAG 371
QY      106 TrpLeuLeuLeuAenProGlnIleMetLeuGlyValHisGlyProGlyPheGlyValSerGlu 125
DB      372 TTTCCT-----AGCCAGTTCAACATGACCTTCCTCCGATCCCATGCTAG 419
QY      126 AppCyLeuLeuTrpLeuAenIleTrpAlaProAlaHisAlaAenThrGlySerLeuAenPro 145
DB      420 GACTCCCTGTACCTTCACTTACACGCGCGCCATACAGAGAGGCTTAACTGCGG 479
QY      146 ValLeuValTrpPheProGlyGlyAlaPheGlySerThrGlySerAlaSerIlePheAenGly 165
DB      480 GTGATGAGTGTGATCCAGCGTGTGCTGCTTTTGGCAAGGCTTCCTGTATGATGT 539
QY      166 SerAlaLeuAlaAlaTrpGlyAenValLeuValValValGlnTrpArgLeuGlyIle 185
DB      540 TCCATGCTGGCTGCTCTGGAGAACGTGTGTGTGTCATTCACAGTACCGCTGGGTGTC 599
QY      186 PheGlyPhePheThrTrpAenProGlnIleAlaProGlyAenTrpAlaPheGlyAenGln 205
DB      600 CTGGGCTTCTTCAAGCATCGAGACAGACACGCAACGCGCAACTGGAGCAACCTCGAACCA 659
QY      206 ValAlaAlaLeuSerTrpValGlnIleAenIleGlnPheGlyGlyAenProSerSer 225
DB      660 GTGCTGTCACTACCTGCTGGCTCAGACAGAAATATCCGCCACTTTGAGGCAACCTCGAACCT 719
QY      226 ValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPro 245
DB      720 GTTCAACATTTTGGCGAGTGTGCGGGTGGACGAGTGTGCTTGTGTGTGTCCTCC 779
QY      246 MetAlaLeuGlyLeuPheHisGlyAlaIleMetGlySerGlyValAlaIleIleProTrp 265
DB      780 ATATCCCAAGAGACTTCCACGAGCATATCGAGAGTGGCGCTTCCCTGCGCGGCG 839
QY      266 LeuGlnAlaHisAenTrpGlyIleValSerGlyAenLeu-----GlnValValAlaHisPhe 283
DB      840 CTATTTGCC-----AGCTCAGCTGATGTCTATCTCCACGAGGTGGGCAACCTG 887
QY      284 CysGlyAenAenAlaSerAenSerGlyAlaLeuLeuArgGlyLeuArgThrIleProSer 303
DB      888 TCTGCCCTGTACCAAGTGTACTGTGAAGCCCTGTGGGCTGCTGGCGGGCAAGAGTAA 947

```

```

QY      304 LysGlyLeuLeuThrLeuSerGlnIleGlySerPheThrArgValValAenGlyVala 323
DB      948 GAGGAGATTTCTTGCATTTAAACAGCTTTCAAGATGATCCCCGAGTGTGATGGGCTC 1007
QY      324 PhePheProAenGlnProLeuAenPheLeuSerGlnIleValAenValAenIleProSer 343
DB      1008 TTCTGCGCCAGGACCCCGAGAGCTGCTGCTGCGCATCTTGCAGCTTGCCTTACG 1067
QY      344 IleIleGlyValAenAenHisGlyGlyPheLeuPheLeuPro-----MetLys 359
DB      1068 ATTGTGCTGTCCACAAATGATTTCCGCTGGCTCATCCCCAAGTCATGAGAGTAT 1127
QY      360 GlnAlaProGlyIleLeuSerGlySerAenLysSerLeuAlaLeuHisLeuIleGlnAen 379
DB      1128 GATACCCAGAAAGAAATGACAGAGAGGCTCCAGGCTGTCTGCAAGAAATGTTAACG 1187
QY      380 IleLeuHisIleProProGlnTrpLeuHisLeuValAlaAenGlyTrpPheHisAenLys 399
DB      1188 CTGCTGATGTGCTCTTCAATTTGTGACTGTGACCTGTGAGGAGATCAATTGGGAGCAAT 1247
QY      400 HisSerLeuThrGlnIleArgAenSerLeuLeuAenPheLeuGlyAenValPhePheVal 419
DB      1248 GGGGATCCCGAGACCTTCCAAAGCGCACTTCAGAGATGATGGCGACTCATGTTGTG 1307
QY      420 ValProAlaLeuIleThrAlaArgTrpHisArgPheAlaGlyAlaProValTrpPheTrp 439
DB      1308 ATCCCTCAGCTCCAAATGACATTTT---CAGTGTTCGCGGCGCCCTGTGTACTTAC 1364
QY      440 GlnPheArgHisArgProGlnCysPheGlnAenTrpThrArgProAlaPheValLysAenPro 459
DB      1365 GAGTTTCAAGATCAAGCCCACTGCTCAAGAAATCAAGGACCGGACATGAAGGAGAC 1424
QY      460 HisAlaAenGlyValArgPheValPheGlyGlyAlaPheLeuGlyAenIleValMet 479
DB      1425 CATGTGATAGGCTTCCTTGTGTTC---AGAAATTCTTGGGCGCAACTGATTA 1481
QY      480 PheGlnGlyAlaThrGlnGlnGlyIleLeuLeuSerArgIleMetLeuGlyTrpAla 499
DB      1482 TTC-----ACTAGAAAGAGAGACGCTTAAAGGAGATGTGAAGTACGTGGGCG 1532
QY      500 ThrPheAlaArgThrGlyAenProAenGlyAenAenPheLeuSerLeuTrpProAlaTrpAen 519
DB      1533 AACTTGGCGAAGATGGAACCCCAATGGCGAGGCTTCCACATGCGCGCTTTCAC 1592
QY      520 LeuThrGlnGlyTrpLeuGlnIleLeuAenMetSerLeuGlyGlnArgLeuGlyGln 539
DB      1593 CAGAGAGAGCATCTGCACTGAACCTTACAGCTGCGGTGGCGGCTCTGAAGGCC 1652
QY      540 ProArgValAenPheTrpThrSerThrIlePro 550
DB      1653 CACAGGCTCAAGTTCTGGAAGAAAGGCGTCC 1685

RESULT 4
US-09-949-016-555
; Sequence 555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 555

```

```

; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-555

Alignment Scores:
Pred. No.: 2,23e-136 Length: 2169
Score: 1230.00 Matches: 248
Percent Similarity: 63.848 Conservative: 91
Best Local Similarity: 46.708 Mismatches: 168
Query Match: 39.958 Indels: 24
DB: 4 Gaps: 8

US-10-023-515-2 (1-581) x US-09-949-016-555 (1-2169)
QY 31 GYProSerAlaGluGlyProGlnArgAnthrArgLeuGlyTyrP1leGlnGlyValGln 50
DB 132 GGCAGAGACTGACCAAGTCCATCCGGACACACACGCGGACAGGTGCTGGGAGTCTT 191
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 192 GTTCATGTGAAGGGGCGCCATGCCGGGCTCAAACTTCTGGGAATTCATTGCGAAG 251
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeu 90
DB 252 CCACCTTAGAGTCCGCTGGATTGTGACCCCTGAGCCCTGGAATCTTGAGTGTGTG 311
QY 91 ArgGlnAlaThrSerTyrProAsnLeuGlyLeuGlnAsn-----SerGlu 105
DB 312 AGGATGAGAACCAACCATCCGGCCATGTGTCTACAGACCCGACCGCAGTGGAGTCAGAG 371
QY 106 TrpLeuLeuLeuAspGlnHisMetLeuValHisTyrProLysPheGlyValSerGlu 125
DB 372 TTTCCTT-----AGCCAGTTCAACATGACCTTCCGATCCGATGCTGAG 419
QY 126 AspGlyLeuTyrLeuAsnL1eTyrAlaProAlaHisAlaAspThrGlySerLysLeuPro 145
DB 420 GACTCCCTGTACTCTGACATCTACACGCGGCGCCATAGCAGTGAAGGCTCTAACTGCGCG 479
QY 146 ValLeuValTyrPheProGlyGlyValAlaPheLysTrpGlySerAlaSerLysPheAspGly 165
DB 480 GTGAAGGTGTGTGATCCACGCTGTGTGTGCTTTTGTGTGATGCTCTTCTGTATGATGCT 539
QY 166 SerAlaLeuAlaAlaTyrGlnAspValLeuValValValGlnTyrArgLeuGlyLe 185
DB 540 TCCATGCTGGCTGCTTGGAGAACGTGGTGTGTATCATCATCAGAACCGCTGGGTGTC 599
QY 186 PheGlyPhePheThrThrTyrAspGlnHisAlaProGlyAsnTyrAlaPheLysAspGln 205
DB 600 CTGGGCTTCTTCAGCAGCTGAGACACAGCACGCGAACCTGGGGCTTACCTGAGACCA 659
QY 206 ValAlaAlaLeuSerTyrValGlnLysAsnL1eGluPhePheGlyGlyAspProSerSer 225
DB 660 GTGGCTGACATCGCTGGGTGTCAGACGAATATCGCCACTTTGGAGGCAACCTGACCGT 719
QY 226 ValThrLysPheGlyGlySerLysAlaLysSerValSerSerLeuLeuLeuSerPro 245
DB 720 GTACACCAATTTTGGGAGCTGGGAGTGGACGAGGTGTCTTCTGCTGTGTGTGCTCC 779
QY 246 MetAlaLysGlyLeuPheHisLysValAlaLysMetGlySerGlyValAlaLysLeuProTyr 265
DB 780 ATATCCCAAGAGACTTTCACCGGACCATCATGAGAGTGGGTGGCTCTCTGCGCGCG 839
QY 266 LeuGlnAlaHisAspTyrGlnLysSerGluAspLeu-----GlnValValAlaHisPhe 283
DB 840 CTCATATGCC-----AGCTCAGCTGATGTCTATCTCCACGCTGGTGGCAACCTG 887
QY 284 CysGlyAlaAsnAlaSerAspSerGlnAlaLeuLeuAspGlyLeuArgThrLysProSer 303
DB 888 TCTGCTGTGACCAAGTATCTTAAGGCTGGTGGCTGCTCTGCGGGCAAGAGTAA 947
QY 304 LysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValAlaAspGlyAla 323

```

```

DB 948 GAGAGATTCTTCGACATTAAACAAGCCTTTCAGATGATCCCGGAGTGTGATGGGCTC 1007
QY 324 PhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAlaLysProSer 343
DB 1008 TTCCTGCCAGGACACCCCGAGAGCTGGGCTTGTCCGACCTTTCAGCTGTCCCTTAC 1067
QY 344 L1eL1eGlyValAsnAsnHisGlyCysGlyPheLeuLeuPro-----MetLys 359
DB 1068 ATTGTGTGTGTCAACAACAATTAATTCGGCTGGCTCATTCGCCAAGTCAATGAGATCTAT 1127
QY 360 GlnAlaProGluL1eLeuSerGlySerAsnLysSerLeuAlaLeuHisLysLysGln 379
DB 1128 GATACCCAGAGGAATATGACAGAGGCTTCCAGGCTGCTCTCAGAAATGTTTAACG 1187
QY 380 L1eLeuHisL1eLeuProGlnTyrLeuHisLeuValAlaAsnGluTyrPheHisAspLys 399
DB 1188 CTGTGATGTGTCTCTCTCAATTTGTGTGATCTGTGAGGAGAGATACATTGGGACAT 1247
QY 400 HisSerLeuThrGluL1eArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 419
DB 1248 GGGGATCCCGACACCTCCAGACCGCAGTCCAGAGATGATGGGCGGACTCCATGTTGTG 1307
QY 420 ValProAlaLeuL1eThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
DB 1308 ATCCCTGACATCCAGTACATTTT---CAGTGTCCCGGGCCCTGTGTACTTCTAC 1364
QY 440 GluPheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAsp 459
DB 1365 GAGTTCACAGCATCAGCCACGCTGGCTCAAGAACATCAGGCCACCGCACATGAAGCAGAC 1424
QY 460 HisAlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspLysMet 479
DB 1425 CATGTGTATGACTTCTTGTGTTC---AGAACTTCTTGTGGGGGCAACTACATTAA 1481
QY 480 PheGluGlyAlaThrGluGluGlyLysLeuLeuSerArgLysMetLysTyrTyrAla 499
DB 1482 TTC-----ACTGAGAGAGAGAGACAGACTTAACAGAGATGATGAAGTACGGGCC 1532
QY 500 ThrPheAlaArgThrGlyAsnProAsnLysAsnAspLeuSerLeuTyrProAlaTyrAsn 519
DB 1533 AACTTGGAGAAATGAGAACCCCAATGCGAGGGCTGCCACACTGGCCGCTGTTCAGC 1592
QY 520 LeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyLysArgLeuLysGlu 539
DB 1593 CAGAGAGACATTACTCTGACGTGAACCTTACAGCCTGGGGTGGCGGGCTTGAAGGCC 1652
QY 540 ProArgValAspPheTyrThrSerThrLysPro 550
DB 1653 CACAGGCTCCAGTTCTGGAAGAAGCGCTGCC 1685

RESULT 5
US-09-595-682B-27
; Sequence 27, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danke, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2191
; TYPE: DNA
; ORGANISM: Homo sapiens

```

US-09-595-682B-27

Alignment Scores:

Pred. No.:	2,276-136	Length:	2191
Score:	1230.00	Matches:	248
Percent Similarity:	63.84%	Conservative:	91
Best Local Similarity:	46.70%	Mismatches:	168
Query Match:	39.95%	Indels:	24
DB:	4	Gaps:	8

US-10-023-515-2 (1-581) x US-09-595-682B-27 (1-2191)

```

QY 31 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlySerGln 50
Db 154 GGCACAGACTCAGCAGCAGTCCCATCCGACACACACGCGGAGAGTCTCGGGAATCTT 213
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
Db 214 GTCCATGTGAAGGGCCCAATGCCGGGATCCAAACCTCTCGGGAATCCATTTGCCAAG 273
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAsnLeu 90
Db 274 CCACCTCTAGTCCGCTCGCATTTGCACCCCTGAGCCCTCCGATCTTGAGTGTGTG 333
QY 91 ArgGluAlaThrSerTyrProAsnLeuGlyLeuGlnAsn-----SerGln 105
Db 334 AAGGAGGAACCAACCATCCGGCATGTGTCTACAGGACCTCACCGCAGTGGAGTCAAG 393
QY 106 TrpLeuLeuLeuAspGlnHisMetLeuValHisTyrProLysPheGlyValSerGln 125
Db 394 TTTCCTT-----AGCCAGTTCAACATGACCTTCTCCGACTCCATGTCTGAG 441
QY 126 AAPGySLeuTyrLeuAsnIleTyrAlaProAlaHisAlaAspThrGlySerLysLeuPro 145
Db 442 GACTGCTGTACTCAGCATCTACACGCCGCCATAGCCATGAGGCTTAACTGCG 501
QY 146 ValLeuValTrpPheProGlyValAlaPheLysThrGlySerAlaSerLysPheAspGly 165
Db 502 GTGATGGTGTGATGCACGCTGTGTGCTGTTGTTTGGCATGGCTTCTTGATATAGT 561
QY 166 SerAlaLeuAlaAlaTyrGlyValAspValLeuValValAlaGlnTyrArgLeuGlyLys 185
Db 562 TCCATGTCTGGCTGCTTGGAGAACCTGGTGTGTCTATCATCTACATCCGCTGAGTCTC 621
QY 186 PheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGln 205
Db 622 CTGGGCTTCTTCAGACATGAGAACAGAACACCGCACTGGGGCTTACTGAGACCA 681
QY 206 ValAlaAlaLeuSerTrpValGlnLysAsnIleGluPheGlyValAspProSerSer 225
Db 682 GTGGCTGCACTACGCTGGGTCCAGACAGAAATTCGCCCATTTGAGAGGCACCTGACCT 741
QY 226 ValThrIlePheGlyGlySerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPro 245
Db 742 GTCCACCAATTTTGGCGAGTCTGGCGGTGACAGAGTGTGCTTCTGCTGTGTGTCTCCC 801
QY 246 MetAlaLysGlyLeuPheHisLysAlaIleMetGlnSerGlyValAlaIleIleProTyr 265
Db 802 ATATCCCAAGACTCTTCCAGCAGACCATCATGAGAGTGGCGGCTCCGCGCCGCGC 861
QY 266 LeuGluAlaHisAspTyrGlyLysSerGlyValAspLeu-----GlnValAlaHisLysPhe 283
Db 862 CTCACTTCC-----AGCTCAGCTGATGTCTATCTCCACGAGTGTGGCCCAACTG 909
QY 284 CysGlyLysAsnAlaSerAspSerGlyAlaLeuLeuArgCysLeuArgThrLysProSer 303
Db 910 TCTGCTGTGACCAAGTTGACTCTGAGGCCCTGTGGCTGCTGCGCGGACAGAGTAA 969
QY 304 LysGluLeuLeuThrLysSerGlnLysThrLysSerPheThrArgValLysAspGlyAla 323
Db 970 GAGAGAGTTTTCGAATTAAACAAGCTTTCAAGATGATCCCCGAGAGTGGAGTGGAGTTC 1029
QY 324 PhePheProAsnGluProLeuAspLeuLeuSerGlnLysAlaPheLysAlaIleProSer 343

```

```

Db 1030 TTCTGCTCCAGGACCCCAAGAGCTGTGGCTGTGCCAATTTCACCTGTCTCCATGC 1089
QY 344 ILeIleGlyValAlaAsnHisGlyCysGlyPheLeuLeuPro-----MetLys 359
Db 1090 ATTGTGGTGTGACAAACATGAAATTTGGCTGTGCTCATCCCAAGTCAATGAGATCTAT 1149
QY 360 GluAlaProGluIleLeuSerGlySerAsnLysSerLeuAlaLeuHisLysIleGlnAsn 379
Db 1150 GATACCCAGAGAAATGACAGAGAGGCTCCAGAGCTGTCTGCAAGAAATGTTAACG 1209
QY 380 ILeLeuHisIleProProGlnTyrLeuHisLysValAlaAsnGlyTyrPheHisAspLys 399
Db 1210 CTGCTGATGTGGCTCTTCAATTGTGACTCTGTGAGGAGAGATCAATTGGGAGCAAT 1269
QY 400 HisSerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 419
Db 1270 GGGGATCCCAAGACCTCTCCAAAGCCAGTTCCAGAGAGATGATGGCGGACTCCATGTGTG 1329
QY 420 ValProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
Db 1330 ATCCCTGACTCCAAATGACATATTT---CAATGTTCCCGGGCCCTGTGTCTTCTAC 1386
QY 440 GluPheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAsp 459
Db 1387 GAGTTTCAAGATCAGCCCAAGCTGTGACATGACATCAATGACATGAGGAGGAGAC 1446
QY 460 HisAlaAspGluValaArgPheValPheGlyValaPheLeuLysGlyAspIleValMet 479
Db 1447 CATGTGATGAGCTTCTTTGTTTTC---AGAAATTTCTTGGGGGCAATCATTA 1503
QY 480 PheGluGlyAlaThrGlyGlyGlyLysLeuLeuSerArgLysMetLysTyrTrpAla 499
Db 1504 TTC-----ACTAGAGAAAGAGAGAGCTTAAAGAGAAATGATGAAGTACTGCGGC 1554
QY 500 ThrPheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpAlaTyrAsn 519
Db 1555 AACTTTCGAGAAATGGAACCCCAATGCGAGAGCTCTGCACACTGCGCCCTGTTCGAC 1614
QY 520 LeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLysGln 539
Db 1615 CAGAGAGAGAAATCTGTGACGTGAACCTACAGCTGTGGGCGGCTGTGAAGGCC 1674
QY 540 ProArgValaAspPheThrTrpSerThrIlePro 550
Db 1675 CACAGCTCAGTTCTGGAAGAGCGCTGCC 1707

RESULT 6
US-09-595-682B-20
/ Sequence 20, Application US/09595682B
/ Patent No. 6800483
/ GENERAL INFORMATION:
/ APPLICANT: Danke, Mary K.
/ APPLICANT: Potter, Philip M.
/ APPLICANT: Houghton, Peter J.
/ TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
/ FILE REFERENCE: SJ-0005
/ CURRENT APPLICATION NUMBER: US/09/595,682B
/ PRIOR FILING DATE: 2000-01-16
/ PRIOR APPLICATION NUMBER: 60/075,258
/ PRIOR FILING DATE: 1998-02-19
/ PRIOR APPLICATION NUMBER: PCT/US99/03171
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 20
/ LENGTH: 1717
/ TYPE: DNA
/ ORGANISM: Oryctolagus cuniculus
US-09-595-682B-20
Alignment Scores:

```

Pred. No.:	7,83e-126	Length:	171
Score:	1440.50	Match:	251
Percent Similarity:	56.92%	Conservative:	78
Best Local Similarity:	43.43%	Mismatch:	182
Query Match:	37.04%	Indels:	67
DB:	4	Gaps:	13

US-10-023-515-2 (1-581) X US-09-595-682B-20 (1-1717)

[illegible]

QY	313	-----	ThyhsSerPheThrAsgValaIAspGly	322
Db	923	CTAGATCTAGTGGCGACCCCAAGAGAACACCCGCTTCTGACCACTGTGATGATGGG	982	
QY	323	AlaPhePheProAsnGluProLeuAspLeuSerGlnIyValaPheValaIlePro	342	
Db	983	GTGCTGTGCGCAAAAGACCTGGACAGATGTTCTGGCAGAGAAATCAACATGCTGCC	1042	
QY	343	SerIleIleGlyValaAsnAsnIleGluCysGlyPheLeuLeuProMet-----	358	
Db	1043	TACATGTGGGAATCAACACAGAGAGATTTGGCTGGATTTCCCAATGCAATGCTGGGC	1102	
QY	359	-----	LySGlnIaPProGlnIleLeuSerGlySer	368
Db	1103	TATCACTCTGTGAAGCGAACTGGACCAAGAGACAGTACAGAACTCTTGGAAGTCC	1162	
QY	369	AsnIySerIleuAlaLeuHleuIleGlnAsnIleuHsiIleProGlnIyLeu	388	
Db	1163	-----TACCATTGTCAATGCTGTAAAGACTACT	1195	
QY	389	HleuValaIalaangIyTyPheHleAspIlyHsiSerLeuThrGlnIleAspSer	408	
Db	1196	CCAGTGGCCACTGAGAGATTTTATAGAGGGCAGATGACCTGTCAAAAGAAAGACTTG	1255	
QY	409	LeuLeuAspLeuLeuGlyAspValaPhePheValaIProAlaLeuIleThrAlaIyTy	428	
Db	1256	TTCTGTGACATGCTTGACAGATTTGTATTGTGGTCCATCTGTGAATGTGGCTGTAC	1315	
QY	429	HisArgAspAlaGlyAlaProValaTyPheTyGlnPheArgHisArgProGlnIyAspPhe	448	
Db	1316	CACAGAGATGCTGGAGCCCCCACTATATGTATAGATATCGGATATCCGCCAAGTTCTCA	1375	
QY	449	GluAspThrIyPProAlaPheValaIyAlaAspHisAlaAspGluValaArgPheValaPhe	468	
Db	1376	TCAGACATGAGACCCCAAGACAGTATAGGGGACCATGAGATGAGATCTTCTGTCTTA	1435	
QY	469	GlyIyAlaPheLeuIyGlyAspIleValaMetPheGluGlyAlaThrGlnIyGlnIyAsp	488	
Db	1436	GGACCCCGTTTTTAAAA-----GAGGTTGCCACAGAAAGAGAAATC	1477	
QY	489	LeuLeuSerArgIyMetMetIyTyTrpAlaThrPheAlaArgThrGlyAspProAsn	508	
Db	1478	AAACTGACCAAGATGCTGATGAATACTGGGCCAACTTGTCTAGGAATGGGAATCCCAAT	1537	
QY	509	GlyAsnAspLeuSerLeuTrpProAlaTyAsnLeuThrGlnIyTyLeuGlnIleAsp	528	
Db	1538	GGAGAAGGCGTCTTCTCAATGCGCAGCATATGACTACAAAGAAAGGTATCTGCAGATTGGA	1597	
QY	529	LeuAsnMetSerLeuGlyIlnaIyGluLeuGlyIProArgValaAspIleTrpThr	546	
Db	1598	GCCACCAACCCAGCGACGCCAGAAACTGAAGACAGAAAGTGGCTTCTTGACT	1651	

RESULT 7

US-09-264-737-3

/ Sequence 3, Application US/09264737A

/ Patent No. 6107549

/ GENERAL INFORMATION:

/ APPLICANT: Feng, Paul C.C.

/ APPLICANT: Ruff, Thomas G.

/ TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via

/ FILE REFERENCE: 38-21(10551) RE3 Pyridine Tolerance

/ CURRENT APPLICATION NUMBER: US/09/264,737A

/ EARLIER FILING DATE: 1999-03-09

/ EARLIER APPLICATION NUMBER: 60/077,377

/ EARLIER FILING DATE: 1998-03-10

/ NUMBER OF SEQ ID NOS: 11

/ SOFTWARE: PatentIn Ver. 2.0

/ SEQ ID NO 3

/ LENGTH: 1701

/ TYPE: DNA

/ ORGANISM: Rabbit

US-09-264-737-3

Alignment Scores:

Pred. No.: 1,34e-125 Length: 1701
 Score: 1138.50 Matches: 250
 Percent Similarity: 56.62% Conservative: 75
 Best Local Similarity: 43.55% Mismatches: 176
 Query Match: 36.98% Indels: 73
 DB: 3 Gaps: 14

US-10-023-515-2 (1-581) x US-09-264-737-3 (1-1701)

QY 12 TTPCyArhePheLeu1LeuGlnProLeu1Leu1YH1aRglnTTPG1YLySer1Y 31
 Db 25 TGGCTCTTC-----CTTGCAGCCCTGCACCCGATGGGTAC-----60
 QY 32 ProSer1aLg1uG1yProGlnaRgVnThArG1eUg1YTr1leGn1yLysGlnVal 51
 Db 61 CCTCCGCGACCA---CCTGTGGTTGACACTGTAAAGGGGAAAGTCTGGGGAGTTCGTC 117
 QY 52 ThVal1eUg1ySerProValProValaRnValPhe1eUg1YValProPhe1a1aPro 71
 Db 118 AGCTGAAGAGATTGGACAGCCCGTGGCCGTCTTCTGGAGATCCCTTGCAGACCC 177
 QY 72 ProLeuG1ySer1eUaRgPheThRnProGlnProAlaSerPro1TrpArVn1eUaRg 91
 Db 178 CCTCTGGATCCCTGAGGTTTGACACACACCTGCAGAAATCATGGAGCCAGTGAAG 237
 QY 92 Glua1aThSer1yProVn1eUyCy1eUaRnSerGlu1TrPhe1eUaRg1n 111
 Db 238 AACACACACCTCCATCCCTCCATGCTCCCGAGAGCC-----GTATCAGG 285
 QY 112 H1aMe1eU-----LysValH1aRyPro---LysPheG1YVal 123
 Db 286 CATATGCTCTCGAGCTTTCACCAAGAAAGAAACATCCCTTAAGTT-----339
 QY 124 SerG1uArPcyLeu1yTrPhe1n1eYrAlaProAlaH1aRTh1ySer1y 143
 Db 340 TCTGAAGACTGCTTACTGTAATTTACACCCCTGCAGCTGACAAAGAGAGCGG 399
 QY 144 LeuProVal1eUVal1TrPheProG1yG1a1aPhe1yThR1ySer1aSer1ePhe 163
 Db 400 CTGCGCGATGATGATGATCATGAGAGTGTGATGATGGTGGAGATCAACCTAT 459
 QY 164 AaRg1ySer1aLeu1a1a1aRyRg1uArVal1eUVal1Val1Gln1yRg1eU 183
 Db 460 GATGACCTGGCTCTTCTGATGAAAGCTGTGGTGAACATTCAGTACCGCTG 519
 QY 184 G1y1ePheG1yPhePheTh1yTrPArGln1a1aProG1yAaR1yAaR1yPhe1y 203
 Db 520 GGCATCTGGGATCTTTCAGACAGAGATGAGACACCGAGGAACTGGGGTCACTTG 579
 QY 204 AaRg1nVal1a1a1eUSe1Tr1yVal1Gln1yAaR1leG1uPhePheG1y1AaRPro 223
 Db 580 GACACAGGCTGCTGGCTGGTGGTCCAGACAAATATGCAATTTGAGGGGACCA 639
 QY 224 SerSerVal1Tr1lePheG1yG1uSer1aG1yAla1eSerVal1Ser1eU1leU 243
 Db 640 GGCCTGTGACCACTTGGAGAGTCAAGAGAGAGTCAAAAGTGTCTATCTTCAATTA 699
 QY 244 Ser1eUaR1a1yG1yLeuPhe1n1eYrAla1eUeG1uSerG1yVal1a1a1e1e 263
 Db 700 TCCCGCTGACCAAGATCTCTTCATGAGCAATTCAGAGTGGCGTGGCCCTCT 759
 QY 264 Pro1y1eUaR1a1a1a1aRyRg1u1ySer1yAaR-----LeuGlnVal1a1a 281
 Db 760 TCCAGTCTC-----TTACAGAAAGACCAAGTCTTGGTGAAGAAATGGC 807
 QY 282 H1ePheCyG1yAaRn1a1aSer1eUaR1a1eUaR1yG1yLeuA1yTr1y 301
 Db 808 ATGAAAGCTGGGTGAACACCACTCGGCTGTATGTTCACTGCCCGCCAGAG 867
 QY 302 ProSer1yG1eU1eU-----Th1eUSe1Gln1y-----312

Db 868 ACAGAGAGAACTGAGGTCATTTGAATGAATTTATGCTCTAGATCTACTT 927
 QY 313 -----Th1ySerPheTh1yRgVal1aRg1yAlaRhePhePro 326
 Db 928 GGCAGCCCAAGAGAACCTGCTTCTGACACCTGATTTGAATGGGTCTGTCGA 987
 QY 327 AaRg1uProLeuAaR1eUSe1Gln1yA1aPhe1yVal1ePheSer1le1eG1y 346
 Db 988 AAAGACCTGCAGAGATCTTGCAGAGAAATTCAAATGCTGCCCTACATGGTGG 1047
 QY 347 ValaRn1a1eG1yCyG1yPhe1eUaRProHe-----358
 Db 1048 ATCAACAGCAAGATTTGGCTGATTTCCCAATGCAATGCTGGGTATCCACTCT 1107
 QY 359 -----LysG1uA1eUaRProG1u1leUSe1Rg1ySer1eU1ySer1eU 372
 Db 1108 GAAGGCAAACTGCACAGAACAGTACAGAACTCTTGTGAAGT-----1155
 QY 373 Ala1eU1a1eU1eGlnaRn1eUaR1eUaR1eU1eU1eU1eU1eU1eU1eU1eU 392
 Db 1156 -----TACCCATTTGCAATGCTCTAAGAGCTGACTCCAGTGGCACT 1200
 QY 393 AaRg1uTrPhe1n1aR1yH1aSer1eUThR1y1eA1aR1yR1ySer1eUaR1eU 412
 Db 1201 GAGAACTTTTGAAGGACAGATGACCTGTCAAAAGAAAGACTTGTCTTGCATG 1260
 QY 413 LeuG1yArVal1PhePheVal1ProAla1eU1eTh1yAlaR1yR1yH1aR1yA1a 432
 Db 1261 CTTCAGATTTGTATTTGGTGTCCATCTGTGAATGTGCTGTACACCAAGATGCT 1320
 QY 433 G1yAlaRProVal1yTrPhe1yRg1uPhe1n1aR1yR1yA1aR1yR1yH1aR1yA1a 452
 Db 1321 GAGCCCCCACTATATGATGTAGTATCGGTATCGCCCAAGCTTTCATCAGACATGA 1380
 QY 453 ProAlaRheVal1yVal1aRn1a1aR1yA1aR1yR1yPhe1yR1yG1yAlaRhe 472
 Db 1381 CCCAAGACAGTGAATGGGACCAATGAGATGATCTCTGTCTTAAGAGCCCCGTTT 1440
 QY 473 Leu1yG1yAaR1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU 492
 Db 1441 TTAATA-----GAGGTGCACAGAAAGAGATCAAACTGAGCAAG 1482
 QY 493 LyMe1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU1eU 512
 Db 1483 ATGTGATGAATACCTGGCCCAACTTCTGATGATGGAAATCCAAATGAGAGGGCTT 1542
 QY 513 Ser1eU1TrProAla1yAaRn1eUThR1yGln1y1eU1eU1eU1eU1eU1eU1eU 532
 Db 1543 CTTCAATGGCCAGCATATGATCAAGAAAGTTACTGTCAGATTTGAGGCCACCCAG 1602
 QY 533 LeuG1yGln1yG1uPhe1yProArGVal1aR1yPhe1yTr1y 546
 Db 1603 GCAGCCCAAGAACTGAAGACAAAGAGATGCTTCTGACT 1644
 RESULT 8
 US-09-799-451-155
 / Sequence 155, Application US/09799451
 / Patent No. 6783969
 / GENERAL INFORMATION:
 / APPLICANT: Tang, Y. Tom
 / APPLICANT: Zhou, Ping
 / APPLICANT: Goodrich, Ryle
 / APPLICANT: Asundi, Vinod
 / APPLICANT: Ren, Feiyan
 / APPLICANT: Zhang, Jie
 / APPLICANT: Xue, Aidong J.
 / APPLICANT: Zhao, Qing A.
 / APPLICANT: Wang, Jian-Rui
 / APPLICANT: Ma, Yundong
 / APPLICANT: Yamazaki, Victoria
 / APPLICANT: Chen, Rui-hong
 / APPLICANT: Wang, Zhimei
 / APPLICANT: Wang, Dunrui


```

; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 155
; LENGTH: 965
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128) .. (925)
US-09-799-451-155

Alignment Scores:
Pred. No.: 5,85e-125 Length: 965
Score: 1129.50 Matches: 222
Percent Similarity: 80.14% Conservative: 0
Best Local Similarity: 80.14% Mismatches: 4
Query Match: 36.68% Indels: 52
DB: 4 Gaps: 1

US-10-023-515-2 (1-581) x US-09-799-451-155 (1-965)
QY 31 GYProSerAlaGlyProGlnArgAsnThrArgLeuGlyTyrPrlIegInGlyLysGln 50
DB 287 GGGCTTCTGCTGAGAGGGCCAGAGAGACACAGAGCTGGATGATTCAGGCGAGACGA 346
QY 51 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 70
DB 347 GTCACTGCTGCTGGAGAGCCCTGCTGCTGGAACGTCTTCTCGAGAGCCCTTTCCTGCT 406
QY 71 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProThrAsnLeu 90
DB 407 CCCCCGCTGGATCCCTGCTGATTTACGAACCCGAGCCCTGATCCGCTGATTAACCTG 466
QY 91 ArgGlnAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGlyTyrLeuLeuAsp 110
DB 467 CGAGAGCACTCTTACCTTAATTTGCTCCAGAACTCAGATGGCTGCTCTTAAGT 526
QY 111 GlnHisMetLeuLysValHisTyrProLysPheGlyValSerGlnAspCysLeuTyrLeu 130
DB 527 CAACGACATGCTCAAGGTGATTAACCCAAATTCGGAGTGTCAAGAGACTGCTTAACCTG 586
QY 131 AsnIleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValIleuValTyrPhe 150
DB 587 AACATCTATGCGCCGCTCCAGCCGATACAGGCTCAAGCTCCCGCTTGGTGTGCTTC 646
QY 151 ProGlyGlyAlaPheLysThrGlySerLysAspTyrIlePheAspGlySerAlaLeuAla 170
DB 647 CAGAGAGGCTCTCAAGACTGCTCAGCTCCTCACTTGTAGTGGGCGCCCTGCTGCTG 706
QY 171 TyrGlnAspValIleuValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 190
DB 707 TATGAGGACGTGCTGTTGTGTGCTCCAGTACCGGCTAGGAATTTGGTTCCTTCA 766
QY 191 ThrTyrAspGlnHisAlaProGlyAsnTyrAlaPheLysAspGlnValAlaAlaLeuSer 210
DB 767 TA----- 768
QY 211 TrpValGlnLysAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGly 230
DB 768 ----- 768
QY 231 GlnSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 250
DB 769 -----CTGCTCCCATGGCCAAAGGCTTA 792
```

```

QY 251 PheHisLysAlaIleMetGlnSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp 270
DB 793 TTCCACAAAGCCATCATGAGAGTGGGTGGCCATCTTCACTGAGGCCCATGAT 852
QY 271 TyrGlnLysSerGlnAspLeuGlnValAlaHisPheCysGlyAsnAsnAlaSerAsp 290
DB 853 TATGAGAGCGTGAAGAGCCTGCAAGCGGTGCAATTTCTGTGTAAACAATCGTCA 912
QY 291 SerGlnAlaLeuLeuArgCysLeuArgThrLysProSerLysGlnLeuLeu 307
DB 913 TCTGAGCCCTGCTGAGGTGCTGAGACAAACCTCCAGAGACTGCTG 963

RESULT 9
US-09-949-016-2515
; Sequence 2515: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2515
; LENGTH: 2802
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2515

Alignment Scores:
Pred. No.: 8.25e-84 Length: 2802
Score: 793.00 Matches: 198
Percent Similarity: 48.00% Conservative: 114
Best Local Similarity: 30.46% Mismatches: 226
Query Match: 25.76% Indels: 112
DB: 4 Gaps: 19

US-10-023-515-2 (1-581) x US-09-949-016-2515 (1-2802)
QY 12 TrpCysPhePheLeuIleLeuGlnProLeuLeuGly-----HisArg 25
DB 256 TGTGCTTGGAGACGACTGCTCTCAATTCCTGAGGACCTGATGTCCTCCCGGAAC 315
QY 26 GlnTyrGlyLysThrGlyPro----- 32
DB 316 ATGTGCTGCGGCTTGGCGCGCCCTGCTGCTCCAGGCCCAAGCCACGCTGGCAGG 375
QY 33 -----SerAlaGln 35
DB 376 AGCTGTGCTTCACCCCTGCTGCTTCTCACTTGGCGCTGAGGCGCAGTACCAAGCCCA 435
QY 36 GlnProGlnArgAsnThrArgLeuGlyTyrIleGlnGlyLysGlnVal----- 51
DB 436 GCACCAACAGTCAACACTCACTTGGAGAGCTAGGGGGTCCCGAGTACCACTGCCACT 495
QY 52 ThrValLeuGlySerProValAsnValPheLeuGlyValProPheAlaIlePro 71
DB 496 GAGATCCTGGGG-----CTGTGACCAATACCTGGGGGTGCTCCCTGACGACTCC 546
QY 72 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeuArg 91
DB 547 CCGATCGCGCGAGAAAGTTCCTGCTGCTCCCTGAAACCAACCCCATCTGATCGGGCATCCGG 606
QY 92 GlnAlaThrSerTyrProAsnLeuCysLeuGlnAsnSerGln----- 105
```

```

Db      607 AACCCACACACTTTCCCGGAGTGTGCCCCGAGAACATCCACAGAGCTGTCCCGAAGTC 666
QY      106 -----TrrpleuLeuAapGlnHishetleuYvAlHsTYrProLYsPhe 121
Db      667 ATGTGCGCGGCTGTTCATCTGACCAACTTGAGTATCGTCTGACTTACATCCAGAG--- 723
QY      122 GlyValSerGluAapCYsLeuTYrLeuAapH1eTYrAlaProAlaHishAlaapThnGly 141
Db      724 -----CCCAACGAAAGACTGTCTCTACTGAAAGTCTATGTGCGAAGAGATGATCCGAGC 780
QY      142 SerLYs----- 143
Db      781 GCTAAGAAACAGCGGAGACTTAGCCGATATATGACGGGAGATGAAGATGAACATCCGG 840
QY      144 -----LeuProValleuValTrrPheProGlyGlyAlaPheYsThnGlySer 159
Db      841 GACAGTGTGCTAAACCCCTCATGTGTCTATCATCCAGAGGCTCTTACATGAAAGGAGACA 900
QY      160 AlASerlePheAapGlySerAlaLeuAlaAlaTYrGluAapValleuValValVal 179
Db      901 GGCACATATGATGTGACAGCATCTTCGCAATTATGGCAATGTCAATCGCATACCCCTC 960
QY      180 GlnTYrArgLeuGlylePheGlyPhePheThrThrTrrAapGlnHishAlaProGlyAan 199
Db      961 AACTATCGGCTTGAAGTGTGAGTATCTTGAGTATCGAGATCAAGGCTCCAAAGGCAAC 1020
QY      200 TrrAlaPheLYsAapGlnValAlaAlaLeuSerTYrValGlnLYsAanH1eGluPhePhe 219
Db      1021 TATGGGCTCTTGAACAGATCCAGGCTTCGCTGGGTAGAGAAATATTCCTTTTC 1080
QY      220 GlyLYsAapProSerSerValThrlePheGlyLYsSerAlGlyAlaH1eSerValSer 239
Db      1081 GGGGAGACCCCGCGGATCACTGTCTTGGCTCGGCAATTGGTGCATCTCGCTCAGC 1140
QY      240 SerleuH1eLeuSerProMetAlaLYsGlyLeuPheH1eLYsAlaH1eMetGluSerGly 259
Db      1141 CTCCTCAAGTGTTCATCATCATCAAGAGGACTTTTCCAGAGGCCATCATCAAAAGTGGC 1200
QY      260 ValAlaH1eIleProTYrLeuGlnAlaH1eAapTYrGluYsSerGluAapLeuGlnVal 279
Db      1201 TCTGTCTGTCCAGCTGG-----GCTGTGACTCAACAACGATGAAGTACACCGCTGG 1254
QY      280 ValAlaH1ePheCYsGlyAanAapH1eSerAapSerGlyAlaLeuLYsAapCYsLeuAap 299
Db      1255 CTGGAGACAAAGTGGCTGTATGTCTGTGACACCGGATATGTGTGAGCTGTCTTCGG 1314
QY      300 ThrlYsProSerLYsGluLeu-----LeuThrlLeuSerGlnLYsThrlYsSer 315
Db      1315 CAAGAAGTGCACAGAGCTGTAGAGACAGACATCCAGCCGCGCTACACAGTGGCC 1374
QY      316 PheThrlArgValAlaAapGlyAlaPhePheProAanGluProLeuAapLeuSerGln 335
Db      1375 TTTGGCCCTGTGATGTATGTATGTATCTCTGATGACCTGTGATCTTCATGAGACAG 1434
QY      336 LYsAlaPheLYsAlaIleProSerleIleGlyValAanAanH1eGlu---CYsGlyPhe 354
Db      1435 GGCAGAGTTCCTCAACATATGATCATGTAGTGTCAACAGGCGAGGGCTCAAGTTT 1494
QY      355 LeuLeuProMetLYsGluAlaProGlnH1eLeuSerGlySerAanLYsSerleuAlaLeu 374
Db      1495 GTGGAAGGGGTGTGAGCCCTGAGAGATGTGTCTCTGCACTGACTTGAATATTCGTC 1554
QY      375 ---H1eIleuH1eGlnAanH1eLeuH1eIleProPro-----GlnTYr 387
Db      1555 TCCAAATTTTGTGACATCTGTATGTGCTATCTTAGGGTAAAGACACCTGCAAGAGCC 1614
QY      388 LeuH1eIleuValAlaAanGluTYrPheH1eAapLYsH1eSerleuThnGluH1eArgAap 407
Db      1615 ATCAAGTTCTATGATACAGACTGG---GCAAGACGTGACAAACCTGAGACCCCGGAAA 1671
QY      408 SerleuLeuAapLeuSerGlyAapValPhePheValValProAlaLeuH1eThrlAanY 427

```

```

Db      1672 ACACTGGTGGCACTCTTCACTGACCAAGTGGGTGAGCCCTCAGTGTGACAGCCGAT 1731
QY      428 TYrH1eArgAapAlaGlyAlaProValTYrPheTYrGluPheAanG1eArgProGlnYs 447
Db      1732 CTGACATGCCCGGTAGGGCTCGGCTCACTTCTTACGCTTCTATCAT-----CACTGC 1785
QY      448 PheGluAapThrlYsProAlaPheValLYsAlaAapH1eAlaAapGluValArgPheVal 467
Db      1786 CAGAGCTTCATGAAGCCCTGTGTGTGATGATGACGTGAGTGGGATGAAGTACCTATGTT 1845
QY      468 PheGlyGlyAlaPheLeuLYsGlyAapH1eValMetPheGlu---GlyAlaThnGlu 486
Db      1846 TTTGGGGTTCCTATGTA---GGCCCCACTGACCTTTTCCCTGCAACTTCTCAAGAAAT 1902
QY      487 GlnLYsLeuLeuSerAapLYsMetMetLYsTYrTrrAlaThrPheAlaArgThGlyAan 506
Db      1903 GATGTATGTCTCAGTGTGCTGTGATGATCTTATGACAACTTTGCAAGAGCTGGGAT 1962
QY      507 ProAan-----GlyAanAapLeuSer 513
Db      1963 CCAACAAAGCCGGTCCCGCAGACCAAGTTCATTCACACCAAGGCCAACCGCTTGAG 2022
QY      514 -----LeuTrrProAlaTYrAanLeuThnGluH1e---TYrleuGlnLeuAapLeuAan 530
Db      2023 GAAGTGGCTGTGTCCAATCAATCCCGAGACAGCTTACTTTCATCATCGGGCTGAAA 2082
QY      531 MetSerleuGlyGlnAapLeuLYsGluProArgValAapPheTrrPthSerThrllePro 550
Db      2083 CCAAGGTTCGAGATCATTAACCGGCGCATTAAGGTGCTTTGAAACATCTGTGGCC 2142
QY      551 LeuIleLeuSerAlaSerAapMetLeuH1e 560
Db      2143 CACCTATACAACTCGCATGACATTTTCAC 2172

RESULT 10
US-09-949-016-1017
; Sequence 1017, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTNER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1017
; LENGTH: 4960
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1017

Alignment Scores:
Pred. No.: 5,996-83
Score: 789.50
Percent Similarity: 51.32%
Best Local Similarity: 33.72%
Query Match: 25.64%
DB: 4
Gaps: 22

US-10-023-515-2 (1-581) x US-09-949-016-1017 (1-4960)
QY      41 ThrlArgLeuGlyTYrPleGlnGly-----LYsGlnValThr-----ValleuGlySer 56
Db      838 ACCACTTGGAAAGATGAAGAGGATTAAGAGAACTCAATATTAATAATTTTGGGG--- 894

```

```

Qy 57 ProValAsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeu 76
Db 895 -----CGTGTATTCAATTTCTTGGGGTTCCATTGACGCCCAACAGGGGAGCGT 948
Qy 77 ArgPheThrAsnProGlnProAlaSerProTyrAspAsnLeuGlyAlaThrSerTyr 96
Db 949 CGTTTACGCCCTCCAGAACACCAACCATCTCCCTGTGCATATCAGAAATGCCCAATTT 1008
Qy 97 ProAsnLeuGlyLeuAsnSerGlnTyrPheLeuLeuAspGlnHisteLeuVal 116
Db 1009 GCTCTGTGTGTCGCCAGAAAT-----ATCATTTGATGCGCAGATTCACAAGATC 1056
Qy 117 HisTyrProLysPhe-----GlyVal 123
Db 1057 ATGCTTCCTGTGTGTTACTAATACTGATGTGGTTTCAATATGTGCACAACGAC 1116
Qy 124 SerGluAspGlyLeuTyrLeuAsnIleTyrAlaProAlaHis-----AlaAspThr 140
Db 1117 AGCGAAGACTGCGCTATTATAATATATATATGTCGCGACTGAGATGATATTCGGGACAGT 1176
Qy 141 GlySerLysLeuProValLeuValTyrPheProGlyGlyAlaPheLysThrGlySerAla 160
Db 1177 GGGGGTCCCAACCGATGATGGTATATCCATGGTGGCTCATATATGGAAGGATCTGGA 1236
Qy 161 SerIlePheAspGlySerAlaLeuAlaTyrGlnAspValLeuValValGln 180
Db 1237 AATTTATATGATGAGAGTCTCTGGCAAGTTATGGCAATGTATGTCTATCACTCAAC 1296
Qy 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTyrAspGlnHisAlaProGlyAsnTyr 200
Db 1297 TATCACTTGGAGTACTGCGATTCTTTGAGTACAGCGCATCAGCGTCAAGGGAACATAT 1356
Qy 201 AlaPheLysAspGlnValAlaAlaLeuSerTyrValGlnLysAsnIleGluPheGly 220
Db 1357 GCACTCCCTTGATCTCATCAAGCTTTAAGATGAGTACAGTGAAGAAATTTCTTTGGT 1416
Qy 221 GlyAspProSerSerValThrIlePheGlyGluSerAlaGlyAlaIleSerValSerSer 240
Db 1417 GGTGACCCCTTAAGATCACTGTTTGGATCTGGTGTGGGGGTTCATGTCTCAACTG 1476
Qy 241 LeuIleLeuSerProMetAla-----LysGlyLeuPhe 251
Db 1477 CTGACTTATCCCATTTATCTGAAGTAAACGTTGAGACAAATTCACAAAGACATTTT 1536
Qy 252 HisValAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAspTyr 271
Db 1537 CAACGAGCAATAGCTCAAGATGGAAACAGCCCTTTCACCTGG-----GCTGTATGATT 1590
Qy 272 GluLysSerGluAspLeuGlnValAlaHisPheCysGlyAsnAspAlaSerAspSer 291
Db 1591 CAACCTGCAAATATGCTAGATGTGGTGGCCACAAAGTTGGTGCATGTTCAATACA 1650
Qy 292 GluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer--- 310
Db 1651 GTAGAGTTAGTGAATGCTACAGAAAGACCTTCAAGAAACCTTGTGACCAAGATATT 1710
Qy 311 GlnLysThrLys-----SerPheThrArgValValAspGlyAlaPhePheProAsn 327
Db 1711 CAACGAGCTCGATACACATACCTTTGGACCTGTGATTTGATGATGATGATTAATACCAAC 1770
Qy 328 GluProLeuAspLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyVal 347
Db 1771 GACCCCGAGATTTATGAGCAGCAGAGAGATTTTCAACTAGATATATATGTTAGAGATG 1830
Qy 348 AsnAsnHisGlu---CysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSer 366
Db 1831 AACCAAGGGGAAGGTTAAATTTGTGTAATAATAGTAAAGATAGGAGATGATATACA 1890
Qy 367 GlySerAsnLysSerLeuAlaLeu---HisLeuIleGlnAsnIleLeuHisIleProPro 385
Db 1891 GCTAGTGAATTTGACTGTGTTTCAAAATTTTGTGATATATGATGATATCTGAA 1950
Qy 386 -----GlnTyrLeuHisLeuValAlaAsnGluTyrPheHisAspLys 399

```

```

Db 1951 GCGAAAGATGTTTGAGAGAAACCATTAAGTTCATGTAATGACTGG-----GCTGACCGT 2007
Qy 400 HisSerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheVal 419
Db 2008 CATAACTCTGAACCGAAGAAAGACATTAAGTGGCTTTGTGTACGACATCAGGGGATG 2067
Qy 420 ValProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyr 439
Db 2068 GCACGAGCTGTAGCCACAGCGGATCTTCACTCAAACTTTGGTTCCTACGATCTTCTAT 2127
Qy 440 GluPheArgHisArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAsp 459
Db 2128 GCCTTTTACAT-----CATTCGCAACAGATCAGTTTCCAGCTTGGGCTGATGCAGCC 2181
Qy 460 HisAlaAspGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMet 479
Db 2182 CACGGAAGCAGAGTCCCTATGTACTGGGAATCCCATGATTT--GGCCCTACAGAGTTA 2238
Qy 480 PheGlu---GlyAlaThrGluGluLysLeuLeuSerArgLysMetMetLysTyrTyr 498
Db 2239 TTTCTTGCATTTCTCCAAATAATGATGTGATGCTGAGTGCAGTTGTATACATATCTGG 2298
Qy 499 AlaThrPheAlaArgThrGlyAsnProAsn-----GlyAsnAspLeuSerLeu--- 514
Db 2239 ACAATTTTGTCTAAACGTGATGACCCAAATCAACACGTCCTCAAGACAGCAAAATTCAT 2358
Qy 515 -----TrpProAlaTyrAsnLeuThrGluGln 523
Db 2359 CATACAAACCCAACTGTTTGAAAGATGACATGAGACAGATATTTCCAGAAACCA 2418
Qy 524 ---TyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluProArgVal 542
Db 2419 CTATTATCTCATATTTGATTAACCAAGATTAAGAACTTACACAGCCCAATAGGATG 2478
Qy 543 AspPheThrThrSerThrIleProLeuIleLeuSerAlaSerAspMet-LeuHisSerPro 562
Db 2479 AACCTGTGTTGAGATGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2538
Qy 562 GluLeuSer-----SerLeuThrPheLeuSerLeuLeu----- 572
Db 2539 TCTACACACATAAAGTCCATCACTCACTGATCACTTTCAGACCTTACGAGAAAAATTC 2598
Qy 573 -----GlnProPhe 575
Db 2599 GTACCTGTACGTCAAGCTTTTC 2620
Qy 2599 GTACCTGTACGTCAAGCTTTTC 2620

```

RESULT 11
 US-09-949-016-2516
 Sequence 2516, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: C1001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 PRIOR FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,766
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2516
 LENGTH: 2862
 TYPE: DNA
 ORGANISM: Human
 US-09-949-016-2516

Alignment Scores:

Pred. No.: 1.78e-82 Length: 2862
 Score: 782.00 Matches: 199
 Percent Similarity: 46.65% Conservative: 114
 Best Local Similarity: 29.66% Mismatches: 224
 Query Match: 25.40% Indels: 134
 DB: 4 Gaps: 20

US-10-023-515-2 (1-581) x US-09-949-016-2516 (1-2862)

```

QY      12  TrpCybRhepHeuLeuGlnProLeuLeuGly-----H1sArg 25
DB      256  TGGTGGCTTGGACAGACCTGCTCTTACATGCTGGACCTGTAGTCCCTCCGCGGAGAC 315
QY      26  GlnTrpGlySerThrGlyPro----- 32
DB      316  ATGAGGTGGGGGCTTGGCCCGCCCTGCTCCCTAGACCCCAAGCCAGGTGGCAAG 375
QY      33  -----Ser1aGlu 35
DB      376  AGCCTGTGCTCAACCTGTGGTTCCTCAGTTTGGCGCTGAGGGCCAGTACCCAGGCCCA 435
QY      36  G1yProGlnArgSerThrArgLeuGlyTrp1LeuGlnGlySerGlnVal----- 51
DB      436  GCACCCACAGTCAACACTACTCTTGGAGAGCTAAGGGGTGCCAGTACCACTGCCAGT 495
QY      52  ThrValLeuGlySerProValProValaenValPheLeuGlyValProPheAla1aPro 71
DB      496  GAGATGCTGGGG-----CCTGTGACCAATACCTGGGGGTGCTTACGCAAGCTGCC 546
QY      72  ProLeuG1SerLeuArgPheRhrAnProGlnProAlaSerProTrpRhePheLeuArg 91
DB      547  CCGATGGGGAGAAACGTTTCCGCCCCCTGAACCAACCCCACTCGGTGCGGCAATCCGG 606
QY      92  G1uAlaThrSerTrpRhePheLeuGlyLeuGlnaenSerGlu----- 105
DB      607  AAGCGCACACACTTTCCTCCAGTGTGCCCGGAAACATCCACACAGCTGTGCCCAAGTC 666
QY      106  -----TrpLeuLeuLeuArgGlnH1sMetLeuValH1sTrpProValPhe 121
DB      667  ATGCTGCCCGCTGTGTTCACTGCACTTGCATTCGATATCGCTACTTACATCCAGAG-- 723
QY      122  G1yValSerGlnaenProGlySerLeuTrpLeuLeuValTrpAlaProAlaH1s----- 137
DB      724  ---CCCAAGAGAGCTGTCTCTCACTGAAAGCTTATGTGCCAGCGAGAGATGAAGCGG 780
QY      137  ----- 137
DB      781  ATTTCCAGGAATGCCCGCGAAGCCCAAGAAATTTGTAGGAAGAGAGATCCGGC 840
QY      138  -----Ala 138
DB      841  GCTAAGAAACAGGGGAGAGACTTACCGGATATGACGGGGATGAAGATGAAGATCCGG 900
QY      139  AspThrGlySerTrpLeuProValLeuValTrpPheProGlyGlyAlaPheLeuThrGly 158
DB      901  GACAGTGGTGTAA--CCCGTATGCTTCAATTCACCGAGGCTCTTACATGAAGGG 957
QY      159  Ser1aSer1ePheRhePheGlySer1aLeuAla1aTrpGluSerValLeuValVal 178
DB      958  ACAGGCAACATGATGATGAGGAGCATCTCCGCAAGTATGGCAATGTCATCGTATCACC 1017
QY      179  ValGlnTrpArgLeuGly1LePheGlyPhePheThrTrpAspGlnH1a1aProGly 198
DB      1018  CTCAACTATCGGGTGGAGTGTAGGTTTCTGTAGTACTGTGAAGTCAAGGCTCCAGGGC 1077
QY      199  AsnTrpAlaPheLeuValaenGlnValAla1aLeuSerTrpValGlnValaen1eGluPhe 218
DB      1078  AACATATGGCTCTTGGACAGATCCAGGCGCTCCGCTGGGGTGAAGCAATATTGCTTC 1137
QY      219  PheGlyGlyAspProSerSerValThr1LePheGlyGluSerAlaGlyAla1eSerVal 238
DB      1138  TTGGGGGAGAACCCCGCGGATCATGTCTTGGCTGGGGCATGGTGTGATCTTCGCTC 1197
  
```

```

QY      239  SerSerLeuLeuSerProMetAlaLeuGlyLeuPheH1sLeuAla1eMetGluSer 258
DB      1198  AGCCTCTCAGCTGTGTCAATCATCAAGAGGACCTTTCCAGAGAGCATCATCCAAAGT 1257
QY      259  G1yValAla1e1e1eProTrpLeuGlnAlaH1sArgTrpGlySerGluAspLeuGln 278
DB      1258  GGCCTGCTGTGTCCAGCTGG-----GCTGTGAACCTCAACCAAGTGAACCAAGCAGC 1311
QY      279  ValValAlaH1sPheCyeGlyAsnAla1aSerAspSerGluAlaLeuLeuArgCyeLeu 298
DB      1312  CTGCTGGCAACAAGGTGGCTGTAAATGTGCTGACACCTGGATATGTGGTACTGTCTT 1371
QY      299  ArgThrTrpProSerTrpGlnLeu-----LeuThrLeuSerGlnTrpTrpLeu 314
DB      1372  CGGCMAAAGAGTGGCCAGAGACCTGTAGAGAGACATCCAGCCAGCCGGTACCAAGTG 1431
QY      315  SerPheThrArgValValaenGlyAlaPhePheProaenGluProLeuAspLeuSer 334
DB      1432  GCCTTTGGCCCTGTGATGTAGTGTAGTGTCAATTCTGTATGACCTTGAGATCTTCATGAG 1491
QY      335  GlnValaPheLeuVala1eProSer1e1eGlyValaenH1sGlu--CyeGly 353
DB      1492  CAGGGCGAGTTCCTCAATATATGATCATGTAGTGTCAACCAAGGGCGAGGCTTCAAG 1551
QY      354  PheLeuLeuProMetTrpGluAlaProGln1LeuSerGlySerAsnTrpSerLeuAla 373
DB      1552  TTTTGTGAAGGGGTGGAGCTGAGACCTGAGAGATGTGTCTGTGGACACTGTTGACTATTCC 1611
QY      374  Leu--H1sLeu1eGlnaen1LeuH1s1e1eProPro-----Gln 386
DB      1612  GTTCCAAATTTTGGACAAATCTGTATGCTATCTGTAGGCTAAGAGACACCTGCCAGAG 1671
QY      387  TyrLeuH1sLeuVala1aenGluTrpPheH1sAspTrpH1sSerLeuThrGluLeuArg 406
DB      1672  ACCATCAAGTTCAATGTATACAGACTGG--GCAGACCTGTACCAACCTGAGACCCCGCT 1728
QY      407  AspSerLeuLeuAspLeuLeuGlyAspValPhePheValProAlaLeu1eTrpAla 426
DB      1729  AAAACATGTGTGGCATCTTTCATCACTGACACCAAGTGGGTGAGGCCCTCAGAGGTGACAGCC 1788
QY      427  ArgTrpH1sArgAspAlaGlyAlaProValTrpPheTrpGluPheArgH1sArgProGln 446
DB      1789  GATCTGCATGCCCGCTACAGGCTGCTGCTCACTTCTACGCGCTTATCAT-----CAC 1842
QY      447  CyePheGluAspRhrTrpProAlaPheValTrpAlaAspH1a1aAspGluValaPhe 466
DB      1843  TGCCAGAGCTCATGAAGCTCTGTGTGTCAGATGCACTCATGGGGATGAAGTACCCAT 1902
QY      467  ValPheGlyGlyAlaPheLeuTrpGlyAsp1LeuAlaMetPheGlu--GlyAlaThrGlu 485
DB      1903  GTTTTGGGGTTCCTATGGTA--GGCCCACTGACCTTTTCCCTGCAACTTCTCCAAAG 1959
QY      486  GlnGluLeuLeuLeuSerTrpGlyMetMetLeuTrpTrpAlaThrPheAlaArgTrpGly 505
DB      1960  AATGATGTTATGCTCAGTGTCTGTCTCATGACCTATGTGACCAACCTTGGCCAAAGCTGGG 2019
QY      506  AsnProAsn-----GlyAsnAspLeu 512
DB      2020  GATCCCAACACACCGGTCCCGCCAGAGACACCAAGTTTCATTCACACCAAGGCCACCGCTTT 2079
QY      513  Ser-----LeuTrpProAlaTrpAsnLeuThrGluGln--TrpLeuGlnLeuAspLeu 529
DB      2080  GAGGAAGTGGCTGTGTCCAAATACATCCCGAGACAGCTACCTTCACTCATCGGGCT 2139
QY      530  AsnMetSerLeuGlyGlnArgLeuLeuGluProArgValaAspPheTrpThrSerThr1e 549
DB      2140  AAACCAAGGCTCCAGATCATTCAGGGCCACTAAGGTGGCTTTTGGAAACATCTGGTG 2199
QY      550  ProLeu1eLeuSer1aSerAspMetLeuH1s 560
DB      2200  CCCCACTATACACCTGCATGACATGTTCCAC 2232
  
```

RESULT 12

US-09-491-356C-5
Sequence 5, Application US/09491356C
Patent No. 6566061
GENERAL INFORMATION:
APPLICANT: Philibert, Robert A.
APPLICANT: Glines, Edward I.
APPLICANT: Delist, Lynn
TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCT4 REGION OF XQ13
FILE REFERENCE: 9465 GUS11
CURRENT APPLICATION NUMBER: US/09/491,356C
CURRENT FILING DATE: 2000-01-26
PRIOR APPLICATION NUMBER: PCT/US99/09365
PRIOR FILING DATE: 1999-04-29
PRIOR APPLICATION NUMBER: 60/083,465
PRIOR FILING DATE: 1998-04-29
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 4233
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4171)..(4171)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4179)..(4179)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4181)..(4181)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4197)..(4197)
OTHER INFORMATION: n is not determined
NAME/KEY: misc_feature
LOCATION: (4228)..(4228)
OTHER INFORMATION: n is not determined
US-09-491-356C-5
Alignment Scores:
Pred. No.: 2,476-81 Length: 4233
Score: 775.00 Matches: 188
Percent Similarity: 50.42% Conservative: 113
Best Local Similarity: 31.49% Mismatches: 220
Query Match: 25.17% Indels: 76
DB: 4 Gaps: 17
US-10-023-515-2 (1-581) x US-09-491-356C-5 (1-4233)
QY 29 LysThrGlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGly 48
DB 506 AGGCGCATGCCAGGAGCCCGGAGCCAGCCAGATCACTCACTTGGGAGGTAAGGGGT 565
QY 49 LysGlnVal-----ThrValLeuGlySerProValProValAsnValPheLeu 64
DB 566 GCCCAGATGACCACTGCCAGATGATCTGGGG-----CTGTGGACCAATACCTGT 616
QY 65 GlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsnProGlnProAla 84
DB 617 GGGGGCCCTACGAGCTCCCGGATCGGCGAGAAAGTTTCTGCCCCCTGAACCAACC 676
QY 85 SerProTrpAspAsnLeuArgGluAlaThrSerTrpProAsnLeuCysLeuGlnAsnSer 104
DB 677 CCATCTGTGGGAGATCGGAGACCACTTTCCCGGAGTGTGCCCCCAAGAACATC 736
QY 105 Glu-----TrpLeuLeuLeuAspGlnIleMetLeu 114
DB 737 CACACAGCTGTGCGGAGATCATGCTGCCGTCTGTCACTGCCAATTGGATATCGTC 796
QY 115 LysValIleTrpProLysPheGlyValSerGluAspCysLeuTrpLeuAsnIleTrpAla 134

DB 797 GCTACTTACATCCAGAG-----CCCAAGAAAGCTCTCTACTCAAGCTATGTG 850
QY 135 ProAlaIleAlaPheThrGlySerLys----- 143
DB 851 CCGACGAGATGATGATCCGGCGCTAAGAAACAGCGGAGACTTAGCGGATATGACGGC 910
QY 144 -----LeuProValIleValTrpPheProGly 152
DB 911 GATGAAGTGAAGACATCCGGAGACGTGCTAAACCCGTCATGTCTTACATCCACGA 970
QY 153 GlyAlaPheLysThrGlySerLysSerIlePheAspGlySerAlaLeuAlaTrpGlu 172
DB 971 GGCTCTTACATGAAGGAGACAGGCAACATGATTGTCAGATCTTGCCGATATGGC 1030
QY 173 AspValIleValValValGlnTrpArgLeuGlyIlePheGlyPhePheThrTrp 192
DB 1031 AATGTCATATGATCACTCACTCACTATGCGGTGAGTGAAGTATGGTTTCTGATCTGGA 1090
QY 193 AspGlnIleAlaProGlyAsnTrpAlaPheLysAspGlnValAlaLeuSerTrpVal 212
DB 1091 GATTCAGCTGCCAAGGAGCACTATGAGCTCTTACCAAGATCCAGCCCTCGCGGTG 1150
QY 213 GlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGlyGluSer 232
DB 1151 AGCGAGATATTCCTCTTCCGGGGAGACCCCGCGGATCATGCTTTGGCTCGGAGC 1210
QY 233 AlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPheIle 252
DB 1211 ATTGCTGATCTGCTGGAGCTCTCTCAAGTTCATCATCTAGAGGACTTTTCCAG 1270
QY 253 LysAlaIleMetGluSerGlyValAlaIleIleProTrpLeuGluAlaIleAspTrpGlu 272
DB 1271 AGAGCATCATCAAGAGTGGCTGTGTCAGCTG-----GCTGTAACTTCAAA 1324
QY 273 LysSerGluAspLeuGlnValValAlaIlePheCysGlyAsnAsnAlaSerAspSerGlu 292
DB 1325 CCAAGTGAAGTACACAGGCTGTGSCAGAACAAAGTGGCTGAATGTCTGSCAGACCGG 1384
QY 293 AlaLeuLeuAspCysLeuAspThrLysPheSerLysGluLeu-----LeuThr 308
DB 1385 GATATGCTGACCTGCTTCCGCAAAAGTGCACAAAGAGCTGTGACAGCAGGACCTTCAG 1444
QY 309 LeuSerGlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGlu 328
DB 1445 CCAAGCCGCTACACAGTGGCTTTGGCCCTGTGATGTAAGGATGATCTTCAATGAC 1504
QY 329 ProLeuAspLeuLeuSerGlnLysAlaPheLysValAlaIleProSerIleIleGlyValAsn 348
DB 1505 CCTGAGATCTCATAGAGACAGGCGGAGTTCTCAACTATGATCATCATGTAGTGTCAAC 1564
QY 349 AsnIleGlu-----CysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGly 367
DB 1565 CAGGCGAGGCTCTCAAGTTTGTGAGAGGCTGTGAGCCCTGAGAGTGTCTCTGCGC 1624
QY 368 SerAsnLysSerLeuAlaLeu-----IleLeuIleGlnAsnIleLeuIleIleProPro 385
DB 1625 ACTGACTTGAATATCTCGCTCCCAATTGTGCAAACTGTGTATGCTGATGCTTAAGGCT 1684
QY 386 -----GlnTrpLeuIleValAlaIleAsnGluTrpPheIleAspLysVal 400
DB 1685 AAGACACCCCTCCAGAGAACATCAAGTTCAATGTATACAGCTG---CGACAGCTGAC 1741
QY 401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuGlyLysValPhePheValVal 420
DB 1742 AACCTGAGACCCCGCGTAAACACTGTGTGACATCTTCACTGACACCAAGTGGTGGAG 1801
QY 421 ProAlaLeuIleThrAlaArgTrpIleArgAspAlaGlyAlaProValTrpPheTrpGlu 440
DB 1802 CCTCAGTGTGACAGCGGATGTGATGCTGCGCTAGCGCTGCTCACTTCTTCAACC 1861
QY 441 PheArgIleAspProGlnCysPheGluAspThrLysPheAlaPheValIleAspIleVal 460

Db 1862 TTCTATCAT-----CAGCCGACAGCCTCATGAAGCCTGGTGCATGACGACTCAT 1915

QY 461 AlaapGluValArgpHeValpHeGlyGlyAlaPheLeuLysGlyAspDLeValMetPhe 480

Db 1916 GGGATGATGAATCCCTATGTTTTTGGGGTTCCTANGNA---GGCCCACTGACCTTTTC 1972

QY 481 Glu---GlyAlaThGluGluGluLysLysLeuLeuSerArgLysMetMetLysTyrTTPala 499

Db 1973 CCCTGCAACTTCTCCAGAAATGATNTATGCTCACTGCTGTCATGACTGACTTATGAGAC 2032

QY 500 ThrPheAlaArgThrGlyLysnProAsn----- 508

Db 2033 AACTTGGCCAGACCTGGGGATGCCAACAAGCCGGTCCCCCGAGACACCAAGTCATTCAAC 2092

QY 509 -----GlyAsnAapLeuSer-----LeuTyrProAlaTyrAsnLeuThrGluGlnA--- 523

Db 2093 ACCAAGGCCAACCCGCTTGAGGAAGTGCCGTGTCGCAAAATTCATATCCCGACGACCACTC 2152

QY 524 TyrLeuGlnLeuAapLeuAsnMetSerLeuGlyGlnArgLeuLysGluProArgValAsp 543

Db 2153 TACCTTCACATCGGACTGAAAACCAAGGCTCGAATCATTAACCGGGCCACTAAGGTGGCC 2212

QY 544 PheTyrTrpSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560

Db 2213 TTTTGAAACATCTGTGTCGCCCACTTATACAACCTGATGACATGATGTTCCAC 2263

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Score:	764.50	Matches:	203
Percent Similarity:	49.83%	Conservative:	97
Best Local Similarity:	33.72%	Mismatches:	9
Query Match:	24.83%	Indels:	82
DB:	1.	Gaps:	22

[illegible]

[illegible][illegible]


```

Db      448  TATGGGGGAGGCTTTACAGATGGGGCTCTCTTACAGCTGATGCAATGAGCGGCTTTCTG 507
Qy      169  ALaAaTtYrGluAaPValLeuValValValGltTtYrAglLeuGlyTlePheGlyPhe 188
Db      508  GTACAGCGCGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 567
Qy      189  PheThr---ThTtYrAaPrglnHlaPrglnHlaPrglnHlaPrglnHlaPrglnHlaPrgln 207
Db      568  CTGGCGCTTCGCGGAGAGCGGAGAGCGGCGGAGAGCGGCGGAGAGCGGCGGAGAGCGGCGG 627
Qy      208  ALaLeuSerTtYrValGlnLysAaenilegIuPhePheGlyGlyAaPrgSerSerValThr 227
Db      628  GCCCTGCACTGGGGGTGAGAGAACTGGGAGCTTGGGGGTGAGAGCGGAGAGCGGAGAGCGG 687
Qy      228  lIepheGlyGluSerPheLaIleSerValSerSerLeuileLeuSerProMetaIa 247
Db      688  CTGTtTGGGAGAGCGGAGAGCGGCGGCGGAGAGCGGCGGAGAGCGGCGGAGAGCGGCGG 747
Qy      248  LyvGlYLeuPheHlaIleMetGluSerGlyValAlaIleIleProTtYrLeuGlu 267
Db      748  CGGGGCTGTTCACAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 807
Qy      268  AlaHlaaPrtYrGluLysSerGluAaPrgLeuGlnValAlaHlaIlePhe-----Cys--- 284
Db      808  GTGGGCAATGGAGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 867
Qy      285  -----GlyAaAaAaIleSerAaPrgSerGlyAlaLeuLeuAaGlyCysLeuAaPrgThrPro 302
Db      868  CCAAGCGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 924
Qy      303  SerLyvGluLeu-----ThIleuSerGlnLysThrLys----- 314
Db      925  GCGGAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 984
Qy      315  SerPheThrAaGlyValAaPrgValAaPhePheProAaenileGluAaPrgLeuLeuSer 334
Db      985  TCCTTCGTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1044
Qy      335  GlnLysAaPheLysValAaIleProSerIleIleGlyValAaAaenileGlyCysGlyPhe 354
Db      1045  GCGGAGAGCTTCCACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1104
Qy      355  LeuLeuProMetLyvGluAlaPrgGlnIleLeuSerGlySerAaPrgSerLeu----- 372
Db      1105  TTTTGTG---GtTtYrCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1158
Qy      373  AlaLeuHlaIleLeuIleGlnAaenileLeuHlaIleProPrglnHlaIleLeuValAla 392
Db      1159  CGGGCGGAGTTCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1218
Qy      393  -----AaenileGlyPheHlaIleSerAaPrgSerLeuThIleuile 405
Db      1219  GAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1275
Qy      406  ArgAaPrgSerLeuLeuAaPrgLeuGlyAaPrgValAaPhePheValValProAlaLeuIleThr 425
Db      1276  AGGGAAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1335
Qy      426  AlaAaPrgTtYrHlaAaPrgAaPrgAlaProValTtYrPheTtYrGluPheAaPrgHlaAaPrg 445
Db      1336  GCTGGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1395
Qy      446  GlnCysAaPheGluAaPrgTtYrProAlaPheValLysAaPrgHlaIleAaPrgValAaPrg 465
Db      1396  TCCAGCGCTC-----TCTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1449
Qy      466  PheValAaPheGlyGlyAlaPheLeuLysGlyAaPrgIleValMetPheGlyGlyAla----- 483
Db      1450  TTTATCTTTGGG-----ATCCCTGTGAGCCCTTCGGAAC 1485
Qy      484  ---ThIleGluGlyLysLeuLeuSerAaPrgLysMetMetLyvTtYrTtYrAlaThrPheAla 502

```

```

Db      1486  TACAGCGAGAGAGAAATCTTCCGCGAGGACTGATGCAATATGGGCGCAATTGGCC 1545
Qy      503  ArgThrGlyAaPrgProAaPrg-----GlyAaAaPrgLeuSerLeuTtYrProAlaTtYrAa 519
Db      1546  CGCAGAGGAGATCCCAATGAGCGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCG 1605
Qy      520  Leu---ThIleGluGlnTtYrLeuGlnLeuAaPrgLeuAaPrgLeuAaPrgLeuAaPrgLeu 537
Db      1606  GCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1665
Qy      538  LyvGluProAaGlyAaPrgPheTtYrSerThrIleProLeuileLeuSerAlaSerAaPrg 557
Db      1666  CGCGGCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1725
Qy      558  MetLeu 559
Db      1726  ACGCTC 1731

RESULT 15
US-09-949-016-1192
; Sequence 1192, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1192
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1192

Alignment Scores:
Pred. No.: 1,358-80 Length: 2158
Score: 764.50 Matches: 203
Percent Similarity: 49.83% Conservative: 97
Best Local Similarity: 33.72% Mismatches: 221
Query Match: 24.83% Indels: 82
DB: 4 Gaps: 22

US-10-023-515-2 (1-581) x US-09-949-016-1192 (1-2158)
Qy      1  MetPrgGlnIleuThrSerSerAlaSerGlnTtYrCysAaPheHlaIleLeuGlnPro 20
Db      135  CTTCGCCCACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 173
Qy      21  LeuLeuGlyHlaAaPrgGlnTtYrGlySerThrGlyProSerAlaGluGlyProGlnAaPrg 40
Db      174  -----GAGAGAGTGGGG----- 185
Qy      41  ThrAaPrgLeuGlyTtYrIleGlnIleLysGlnValThrValLeuGlySerProVal----- 58
Db      186  CTGAGGCGCGGAGATGAGAGCTG-CTGCTGACGCTGCGTGGGGCGGCGGCGGCGGCGGCGG 244
Qy      59  -----ProValaAaPrgLeuGlyValProPheAla 70
Db      245  ATTCGCTTGAAGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 304
Qy      71  ProProLeuGlySerIleuAaPrgPheThrAaPrgProGlnProAlaSerProTtYrAaPrgLeu 90
Db      305  CACACCATGAGACCCCGTGTCTTTCTGCGACCGGAGCCCAAGACGACTTGTGTGAGGGGTG 364

```

```

Qy 91 ArgGluAlaThrSerTyrProAsnLeuCyLeuGlnAsnSerGluTrpLeuLeuAsp 110
Db 365 GTAGACGCTACACCTTCCAGAGTGTCTGTAACAATATGTGGACACCTTATCCAGCT 424
Qy 111 GlnH1MetLeuLysValHisTyrProLysPheGlyValSerGluAspCyLeuTyrLeu 130
Db 425 TTTGAGGCGACCGAGATGTGGAAACCCGATGAGCGAGAGACATGCTGTATACCTC 484
Qy 131 AsnLeuTyrAlaProAlaHisAlaAspHisLeuSerLysLeuProValLeuValTrpHe 150
Db 485 AACGGTGGACACCATACCCCGGCTTACA--TCCCCACCCCTGTCTGTCTGTGATC 541
Qy 151 ProGlyGlyAlaPheLysThrGlySerAlaSer-----LlePheAspGlySerAlaLeu 168
Db 542 TATGGGGTGGCTTCTACAGTGGGGCTCTCTTGGACGTGATAGATGGCCGCTTGTG 601
Qy 169 AlaAlaTyrGluAspValLeuValValValGlnTyrAlaGluGlyLlePheGlyPhe 188
Db 602 GTACAGCGCCGAGAGGCTGTGCTGTGCATGAATCCGCGGGGGAGCCCTTGGCTTC 661
Qy 189 PheThr---ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGluValAla 207
Db 662 CTGCGCCCTGCGCGGAGCCGAGAGGCCCGGAGCATGTGGCTCTCTGATTCAGAGCTG 721
Qy 208 AlaLeuSerTrpValGlnLysAsnLleGluPhePheGlyGlyAspProSerSerValThr 227
Db 722 GCCCTGCAGTGGGTGCAGAGAACCTGGCACGCTTCCGGGGGTGACCCGACATCACTGAC 781
Qy 228 LlePheGlyGluSerAlaGlyAlaLleSerValSerSerLeuLleLeuSerProMetAla 247
Db 782 CTGTTTGGGAGAGCGCGGAGCCGCTCGGTGGGCATGCACCTGTGCCCCGCCAGC 841
Qy 248 LysGluLeuPheHisLysValAlaLleMetGluSerGlyValAlaLleProTyrLeuGlu 267
Db 842 CGGGGCTGTTCACAGGGCGGTGCTGCAGAGCGGTGCCCAATGACCCCTGGCCGACG 901
Qy 268 AlaHisAspTyrGluLysSerGluAspLeuGlnValAlaHisPhe-----Cys--- 284
Db 902 GTGGGCAATGGAGAGAGCCCGTCGACGGGCCACGACGCTGCCACCTTGTGGCTGTCT 961
Qy 285 -----GlyAsnAsnAlaSerAspSerGluAlaLeuLysArgCyLeuAlaTyrThrLysPro 302
Db 962 CCAGGCGGCACTGGTGGATGACACAGAG---CTGTAGGCTGTGCTTCCGACACGACCA 1018
Qy 303 SerLysGluLeuLeu-----ThrLeuSerGlnLysThrLys----- 314
Db 1019 GCGCAGGCTCTGTGTAAACCAAGATGGCACTGTGCTTCAAGAAAGCTTCCGAGTTC 1078
Qy 315 SerPheThrArgValAlaAspGlyAlaPhePheProAsnGluProLeuAspLeuSer 334
Db 1079 TCTTTCTGCTGTGTGTATGTGAATCTTCTCACTGACACCCGAGAGCCCTCATCAAC 1138
Qy 335 GlnLysAlaPheLysValAlaLleProSerLleLleGlyValAsnAsnHisGluCysGlyPhe 354
Db 1139 GCGGAGAGACTTCCAGGCTGAGGTGCTGGTGGGTGTGTGAAGAGATGAGGGCTGTAT 1198
Qy 355 LeuLeuProMetLysGluAlaProGluLleLeuSerGlySerAsnLysSerLeu----- 372
Db 1199 TTTCTG---GTTTACGGGGCCCAAGC---TTCAGCAAGACACAGAGTCTTCTATCAGC 1252
Qy 373 AlaLeuHisLeuLleGlnAsnLleLeuHisLleProProGlnTyrLeuHisLleuValAla 392
Db 1253 CCGGCGCAAGTTCCTGGCGGGGTGCGGGTTCCTCCAGGTAAAGTGAACCTGGCAGCC 1312
Qy 393 -----AsnGluTyrPheHisAspLysLysSerLeuThrGluLle 405
Db 1313 GAGGCTGTGTCTTCATTAACAGACTGGGTGATCCGAGAGACCG---GCAAGCTGTG 1369
Qy 406 ArgAspSerLeuLeuAspLeuGluGlyAspValPhePheValValProAlaLeuLleThr 425
Db 1370 AGGAGCGCTGAGAGATGTGGTGGCGACCAATGTCTGTGTGCCCTGGGCCAGCTGTG 1429
Qy 426 AlaArgTyrHisArgAspAlaGlyAlaProValTyrPheTyrGluPheArgHisArgPro 445

```

```

Db 1430 GCTGGGCGACTGGCTGGCCCAAGGTTGCCCGGGGTCTACAGCTGCTTGTGAACACCGTCT 1489
Qy 446 GlnCysPheGluAspTrpHisLysProAlaPheValLysAlaAspHisAlaAspGluValArg 465
Db 1490 TCCAGCGCTC-----TCTGTGCCCTGTGTGATGTGGGGGTGCCCAAGGCTTACAGATCGAG 1543
Qy 466 PheValPheGlyGlyAlaPheLeuLysGlyAspLleValMetPheGluGlyAla----- 483
Db 1544 TTCATCTTTGGG-----ATCCCGCTGACCCCTTCGAAAC 1579
Qy 484 ---ThrGluGluGluLysLeuLeuSerArgLysMetMetLysTyrTrpAlaThrPheAla 502
Db 1580 TACACGGCAGAGAGAAATCTTCCGACGAGCATGTATCGAGTGTGGCCCAACTTGTGCC 1639
Qy 503 ArgThrGluAsnProAsn-----GlyAsnAspLeuSerLeuTrpProAlaTyrAsn 519
Db 1640 CGCACAGGGGATCCCAATGAGCCCCGAGACCCCAAGGCCCAATGAGCCCCGTAACAG 1699
Qy 520 Leu---ThrGluGlnTyrLeuGlnLysAspLeuAsn---MetSerLeuGlyGlnArgLeu 537
Db 1700 GCGGGGCTCAGCAGTACGTTAGTCTGGACCTGGCGCGCTGAGAGCTGGCGGGGCTG 1759
Qy 538 LysGluProArgValAspPheTrpThrSerThrLleProLeuLleLeuSerAlaSerAsp 557
Db 1760 CGCGCCGAGGCTGCGCTTGTGAACCGCTTCTCCCAATATGCTCAGCGCCACCGAC 1819
Qy 558 MetLeu 559
Db 1820 ACGCTC 1825

```

Search completed: June 15, 2005, 13:34:22
Job time : 331 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 22:52:56 ; Search time 1388.55 Seconds

(without alignment)
9634.209 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 2158

Sequence: 1 ccaacgcctccgaataacag5.....aaaaaaaaaaaaaaaaa 2158

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158	100.0	2158	US-10-023-515-1	Sequence 1, Appli
2	2158	100.0	2158	US-10-674-636-1	Sequence 15, Appli
3	2158	100.0	2158	US-10-757-282-45	Sequence 45, Appli
4	1953	96.5	2232	US-10-433-256-23	Sequence 23, Appli
5	1917	88.8	2092	US-10-104-047-249	Sequence 249, Appli
6	1746	80.9	1746	US-10-023-515-3	Sequence 3, Appli
7	1746	80.9	1746	US-10-674-636-3	Sequence 3, Appli

8	1652.8	76.6	1728	US-10-451-168-40	Sequence 40, Appli
9	1579.4	73.2	1962	US-10-094-749-736	Sequence 736, App
10	1435.6	66.5	1746	US-10-114-270-195	Sequence 195, App
11	1432.4	66.4	1746	US-10-451-168-41	Sequence 41, Appli
12	1386.2	64.2	1857	US-10-381-898-20	Sequence 20, Appli
13	1156	53.6	2145	US-10-233-933A-3	Sequence 3, Appli
14	1063.2	49.3	1629	US-10-233-933A-1	Sequence 1, Appli
15	1021.4	47.3	1071	US-10-451-168-42	Sequence 42, Appli
16	508.6	23.6	965	US-10-302-172-155	Sequence 155, App
17	406.6	18.8	2191	US-09-880-107-3854	Sequence 3854, Ap
18	406.6	18.8	2191	US-09-954-531-1038	Sequence 1038, Ap
19	406.6	18.8	2191	US-09-873-367C-828	Sequence 828, App
20	406.6	18.8	2191	US-09-873-367C-829	Sequence 829, App
21	406.6	18.8	2191	US-10-858-771-27	Sequence 27, Appli
22	406.6	18.8	2191	US-10-843-641A-828	Sequence 828, App
23	406.6	18.8	2191	US-10-843-641A-829	Sequence 829, App
24	406.6	18.8	2191	US-10-843-641A-2105	Sequence 2105, Ap
25	406.6	18.8	2484	US-09-925-258-271	Sequence 271, App
26	406.6	18.8	2484	US-10-102-806-271	Sequence 271, App
27	400.2	18.5	2092	US-10-956-157-2153	Sequence 4153, Ap
28	400.2	18.5	3824	US-09-931-836-22	Sequence 22, Appli
29	400.2	18.5	3824	US-10-036-342-22	Sequence 22, Appli
30	400.2	18.5	3824	US-10-036-041-22	Sequence 22, Appli
31	400.2	18.5	3824	US-10-028-072-541	Sequence 541, App
32	400.2	18.5	3824	US-10-035-855-22	Sequence 22, Appli
33	400.2	18.5	3824	US-10-140-808-541	Sequence 541, App
34	400.2	18.5	3824	US-10-121-049-541	Sequence 541, App
35	400.2	18.5	3824	US-10-123-904-541	Sequence 541, App
36	400.2	18.5	3824	US-10-140-470-541	Sequence 541, App
37	400.2	18.5	3824	US-10-175-746-541	Sequence 541, App
38	400.2	18.5	3824	US-10-176-921-541	Sequence 541, App
39	400.2	18.5	3824	US-10-176-921-541	Sequence 541, App
40	400.2	18.5	3824	US-10-227-884-209	Sequence 209, App
41	400.2	18.5	3824	US-10-036-214-22	Sequence 22, Appli
42	400.2	18.5	3824	US-10-137-865-541	Sequence 541, App
43	400.2	18.5	3824	US-10-140-474-541	Sequence 541, App
44	400.2	18.5	3824	US-10-035-719-22	Sequence 22, Appli
45	400.2	18.5	3824	US-10-142-431-541	Sequence 541, App

ALIGNMENTS

RESULT 1
US-10-023-515-1
; Sequence 1, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A. J.
; TITLE OF INVENTION: 5310, A NOVEL HUMAN CARBOXYESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT FILING DATE: 2001-12-18
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-023-515-1

Query Match 100.0%; Score 2158; DB 13; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGCGTCCGAAAAACAGGCGTGGAGACGATGTGAGTAAAGCATGTATATAAAAA 60
Db 1 CCAAGCGTCCGAAAAACAGGCGTGGAGACGATGTGAGTAAAGCATGTATATAAAAA 60
QY 61 TTTAAAAATTATCTTAAATAAAGTAAAGTAAATCCCAATGCCACAGGGACTTAATCTTAC 120
Db 61 TTTAAAAATTATCTTAAATAAAGTAAAGTAAATCCCAATGCCACAGGGACTTAATCTTAC 120
QY 121 CTTCAAGATGCTGCTTTTCTGATTTCTCAAGCCCTGTTGGGACACAGCAAGTGGGAA 180
Db 121 CTTCAAGATGCTGCTTTTCTGATTTCTCAAGCCCTGTTGGGACACAGCAAGTGGGAA 180
QY 181 AAATGGGCGCTTCTGCTGAAGGGCACAGAGAAACAAGGCGTGGATGATTCAGAGCA 240
Db 181 AAATGGGCGCTTCTGCTGAAGGGCACAGAGAAACAAGGCGTGGATGATTCAGAGCA 240
QY 241 AGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAAGCTGTCCTCGAGTCCCTTTG 300
Db 241 AGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAAGCTGTCCTCGAGTCCCTTTG 300
QY 301 CTGCTCCCCGCTGGGATCCCTGGCATTTTACGAAACCGGACGCTGATGCCCTGGGATA 360
Db 301 CTGCTCCCCGCTGGGATCCCTGGCATTTTACGAAACCGGACGCTGATGCCCTGGGATA 360
QY 361 ACTTGGGAGAGCACTCTCTACCTTAATTTGCTCTCCAGAACTCAGAGTGGCTGCT 420
Db 361 ACTTGGGAGAGCACTCTCTACCTTAATTTGCTCTCCAGAACTCAGAGTGGCTGCT 420
QY 421 TTGATCAACAATGCTCAAGGTGTCATTAACCGAAATTCGAGTGTCAAGAGCTGCT 480
Db 421 TTGATCAACAATGCTCAAGGTGTCATTAACCGAAATTCGAGTGTCAAGAGCTGCT 480
QY 481 ACCTGAACATCTATGCGCTGCGCAAGCGGATACAGGCTCCCGCTGGTGGT 540
Db 481 ACCTGAACATCTATGCGCTGCGCAAGCGGATACAGGCTCCCGCTGGTGGT 540
QY 541 GGTTCGCCAGAGGTGCTTCAAGACTGCTCAGCTCATCTTTATGAGTCCGCGCTGG 600
Db 541 GGTTCGCCAGAGGTGCTTCAAGACTGCTCAGCTCATCTTTATGAGTCCGCGCTGG 600
QY 601 CTGCTTAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 CTGCTTAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 TCACCAATGAGATCAGATGCTCGGGGAACTGGGCTTCAAGGACAGAGTGGCTGCT 720
Db 661 TCACCAATGAGATCAGATGCTCGGGGAACTGGGCTTCAAGGACAGAGTGGCTGCT 720
QY 721 TGTCTGGGTCCAGAAAGCATTCAGATTTCTCGTGGGACCCCAAGCTGCTGACATCT 780
Db 721 TGTCTGGGTCCAGAAAGCATTCAGATTTCTCGTGGGACCCCAAGCTGCTGACATCT 780
QY 781 TTGGCAGTCCGCGGAGCCATTAAGTGTCTTCACTTATCTGCTCCCAAGGCAAG 840
Db 781 TTGGCAGTCCGCGGAGCCATTAAGTGTCTTCACTTATCTGCTCCCAAGGCAAG 840
QY 841 GCTTATTCACAAAGCCATTCAGAGTGGGCTGAGCAAAACCTCCAGAGACTGCTGACC 900
Db 841 GCTTATTCACAAAGCCATTCAGAGTGGGCTGAGCAAAACCTCCAGAGACTGCTGACC 900
QY 901 ATGATTAAGAAAGTGAAGACTGAGAGTGTGACATTTCTGTGTTAAACAATGCT 960
Db 901 ATGATTAAGAAAGTGAAGACTGAGAGTGTGACATTTCTGTGTTAAACAATGCT 960
QY 961 CAGACTTGAAGGCTGCTGAGGCTGAGCAAAACCTCCAGAGACTGCTGACC 1020
Db 961 CAGACTTGAAGGCTGCTGAGGCTGAGCAAAACCTCCAGAGACTGCTGACC 1020
QY 1021 TCAGCCAGAAACAAATCTTCACTGAGTGTGATGCTGCTTTCTTCTTAATGAGC 1080
Db 1021 TCAGCCAGAAACAAATCTTCACTGAGTGTGATGCTGCTTTCTTCTTAATGAGC 1080

QY 1081 CTCTAGATCTATTTGTCTGAGAAAGCATTTAAAGCAATTCCTTCCATCATCGGAGTCA 1140
Db 1081 CTCTAGATCTATTTGTCTGAGAAAGCATTTAAAGCAATTCCTTCCATCATCGGAGTCA 1140
QY 1141 ACCAGAGTGTGGCTTCTGCTGCTGCTTAAAGAGGCTCTGAGATCTCAGTGGCTCA 1200
Db 1141 ACCAGAGTGTGGCTTCTGCTGCTGCTTAAAGAGGCTCTGAGATCTCAGTGGCTCA 1200
QY 1201 ACAAGTCCCTTGCCCTCCATCATGATACAAACATCTCGACATCCCGCTCAGATTTGC 1260
Db 1201 ACAAGTCCCTTGCCCTCCATCATGATACAAACATCTCGACATCCCGCTCAGATTTGC 1260
QY 1261 ACCTTGGCTTAATGATTAATCTTCAATGCAAGCACTCCCTGACTGAATCCGAGACATG 1320
Db 1261 ACCTTGGCTTAATGATTAATCTTCAATGCAAGCACTCCCTGACTGAATCCGAGACATG 1320
QY 1321 TTCTGGAATCTTGGAGATGTTCTTTGTGTGCTGCTGCACTGATACAGCTCGATATC 1380
Db 1321 TTCTGGAATCTTGGAGATGTTCTTTGTGTGCTGCTGCACTGATACAGCTCGATATC 1380
QY 1381 ACAAGATGCTGTGACACTGTCTACTTCTATGAGTTCCGACACCGGCTCAGTGGCTTG 1440
Db 1381 ACAAGATGCTGTGACACTGTCTACTTCTATGAGTTCCGACACCGGCTCAGTGGCTTG 1440
QY 1441 AAGACAGAAAGCCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTGTTCG 1500
Db 1441 AAGACAGAAAGCCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTGTTCG 1500
QY 1501 GTGTGCTTCTGTAAGGGGAGCATTTGTATGTTCCAGAGAGCGAGAGAGAGAGT 1560
Db 1501 GTGTGCTTCTGTAAGGGGAGCATTTGTATGTTCCAGAGAGCGAGAGAGAGAGT 1560
QY 1561 TACTGAGCCGAGATGATGAATATCTGGGCTACCTTTGCTCAACCGGGAATCTTAATG 1620
Db 1561 TACTGAGCCGAGATGATGAATATCTGGGCTACCTTTGCTCAACCGGGAATCTTAATG 1620
QY 1621 GGAAGCACTGCTCTGTGGCCAGCTTATATCTGACTGAGAGTACCTTCACTGACT 1680
Db 1621 GGAAGCACTGCTCTGTGGCCAGCTTATATCTGACTGAGAGTACCTTCACTGACT 1680
QY 1681 TGAACATGAGCTCGGACAGAGACTCAAGAACCGGGGTGATTTTGGACAGACCA 1740
Db 1681 TGAACATGAGCTCGGACAGAGACTCAAGAACCGGGGTGATTTTGGACAGACCA 1740
QY 1741 TCCCGCTGATCCGTGCTGCTCGACATGCTCCAGATCTCTTCTTAACTTTC 1800
Db 1741 TCCCGCTGATCCGTGCTGCTCGACATGCTCCAGATCTCTTCTTAACTTTC 1800
QY 1801 TCTCTCTCCAGCCCTTCTTTTCTTTGTGCTCTTGAAGATCTTCTTCTGAT 1860
Db 1801 TCTCTCTCCAGCCCTTCTTTTCTTTGTGCTCTTGAAGATCTTCTTCTGAT 1860
QY 1861 TTTGGTTTCCCTTCTCTCCATTAATTTCTCCGCAATCAATAGCTTCTTCTGAGCTCA 1920
Db 1861 TTTGGTTTCCCTTCTCTCCATTAATTTCTCCGCAATCAATAGCTTCTTCTGAGCTCA 1920
QY 1921 GCTGCTTCTATGGGATCTTGAAGAAACAGCTGCTTCCGTATTTTATGAGACTTA 1980
Db 1921 GCTGCTTCTATGGGATCTTGAAGAAACAGCTGCTTCCGTATTTTATGAGACTTA 1980
QY 1981 GGAATGATCTTACAGAAATCTTTCACATCAAAAGTGCATTTGTCTTGAAGGCA 2040
Db 1981 GGAATGATCTTACAGAAATCTTTCACATCAAAAGTGCATTTGTCTTGAAGGCA 2040
QY 2041 CAAGATTTCTTAATAAATTTGGAAGAGGCTGCTTATGTTGTCTTAATATGCTTT 2100
Db 2041 CAAGATTTCTTAATAAATTTGGAAGAGGCTGCTTATGTTGTCTTAATATGCTTT 2100
QY 2101 TGTAACTCATATGAATAAATCAAGATGTAAATAGGAAAAAATTTTTTTTTT 2158
Db 2101 TGTAACTCATATGAATAAATCAAGATGTAAATAGGAAAAAATTTTTTTTTT 2158

RESULT 2
US-10-674-636-1
; Sequence 1, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A. J.
; TITLE OF INVENTION: S100s-Santlago, Immaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674,636
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-674-636-1

Query Match 100.0%; Score 2158; DB 18; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAAGCGTCCGAAAAACAGGCTTGAGAGCAATGTGAGTAAGCAATGTAAATAAACA 60
DB 1 CCACCGGCTCCGAAAAACAGGCTTGAGAGCAATGTGAGTAAGCAATGTAAATAAACA 60
QY 61 TTTAAATATTTCTTAATTAATAAGTACGAATCCGAATGCCAGAGGACTTACTTCACTG 120
DB 61 TTTAAATATTTCTTAATTAATAAGTACGAATCCGAATGCCAGAGGACTTACTTCACTG 120
QY 121 CTTCAAAATGCTCTTTTCTGATTTCTCAGAGCCCTGTGGAGACAGACAGTGGGAA 180
DB 121 CTTCAAAATGCTCTTTTCTGATTTCTCAGAGCCCTGTGGAGACAGACAGTGGGAA 180
QY 121 CTTCAAAATGCTCTTTTCTGATTTCTCAGAGCCCTGTGGAGACAGACAGTGGGAA 180
DB 121 CTTCAAAATGCTCTTTTCTGATTTCTCAGAGCCCTGTGGAGACAGACAGTGGGAA 180
QY 181 AAACCTGGGCTTCTGCTGAGGAGCCACAGAGAACACAGGCTGGAGTGAATTCAGGGCA 240
DB 181 AAACCTGGGCTTCTGCTGAGGAGCCACAGAGAACACAGGCTGGAGTGAATTCAGGGCA 240
QY 241 AGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTAAGCTGTTCTCGAGTCCCTTTG 300
DB 241 AGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTAAGCTGTTCTCGAGTCCCTTTG 300
QY 301 CTGCTCCCCCGTGGAGATCCCTGCGATTTACGAAACCCGAGCCTGCATCGCCCTGGATA 360
DB 301 CTGCTCCCCCGTGGAGATCCCTGCGATTTACGAAACCCGAGCCTGCATCGCCCTGGATA 360
QY 361 ACTTGGGAGAACCTCTCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTCT 420
DB 361 ACTTGGGAGAACCTCTCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTCT 420
QY 421 TAGATCAACACATGCTCAAGGTCATTAACCGAATTTGGAGTGTCAAGAGCTGGCTCT 480
DB 421 TAGATCAACACATGCTCAAGGTCATTAACCGAATTTGGAGTGTCAAGAGCTGGCTCT 480
QY 481 ACCTGAACATCTATGCGCTGCGACAGCGGATACAGGCTCCAGCTCCCGCTCTGGTGT 540
DB 481 ACCTGAACATCTATGCGCTGCGACAGCGGATACAGGCTCCAGCTCCCGCTCTGGTGT 540
QY 541 GGTTCACAGAGAGTCTTCAAGATGCTGCTCACTCTTTATGTGGTTCGCGCTGG 600
DB 541 GGTTCACAGAGAGTCTTCAAGATGCTGCTCACTCTTTATGTGGTTCGCGCTGG 600

QY 601 CTGCTATAGAGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 601 CTGCTATAGAGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 661 TCACCAATGGGATCAGCATGTCTCCGGGAACTGGGCTTTCAAGAACAGATGGCTGCTC 720
DB 661 TCACCAATGGGATCAGCATGTCTCCGGGAACTGGGCTTTCAAGAACAGATGGCTGCTC 720
QY 721 TGTCTGGGATCCAGAAACATCGAGTTCTTCGGTGGGAGCCAGCTCTGTGACATCT 780
DB 721 TGTCTGGGATCCAGAAACATCGAGTTCTTCGGTGGGAGCCAGCTCTGTGACATCT 780
QY 781 TTGGGAGTCCGCGGAGCCATTAAGTGTCTTATGCTTATATCTCTCCATGGCCAAAG 840
DB 781 TTGGGAGTCCGCGGAGCCATTAAGTGTCTTATGCTTATATCTCTCCATGGCCAAAG 840
QY 841 GGTATTTCCACAAAGCCATCATGAGATGGGATGGCCATCATCCCTTAACTGGAGGCC 900
DB 841 GGTATTTCCACAAAGCCATCATGAGATGGGATGGCCATCATCCCTTAACTGGAGGCC 900
QY 901 ATGATTATGAGAGAGTGAAGGACCTGAGAGTGGTTGCAATTCTGTGTATCAATGCGT 960
DB 901 ATGATTATGAGAGAGTGAAGGACCTGAGAGTGGTTGCAATTCTGTGTATCAATGCGT 960
QY 961 CAGACTTGAAGCCCTGCTGAGAGTGTCTGAGAGCAAAAACCTTCAAGAGCTGTGACC 1020
DB 961 CAGACTTGAAGCCCTGCTGAGAGTGTCTGAGAGCAAAAACCTTCAAGAGCTGTGACC 1020
QY 1021 TCAGCCAGAAAAAAGTCTTTCACTCGAGTGTGTGATGTGTCTTTCTTAAATGAGC 1080
DB 1021 TCAGCCAGAAAAAAGTCTTTCACTCGAGTGTGTGATGTGTCTTTCTTAAATGAGC 1080
QY 1081 CTCTGATCTATTTGTCTGAGAAAGCATTTAAAGCAATTCCTTCATCCGAGTCAATA 1140
DB 1081 CTCTGATCTATTTGTCTGAGAAAGCATTTAAAGCAATTCCTTCATCCGAGTCAATA 1140
QY 1141 ACCAGAGTGTGCTTCTGCTGCTCTTATGAGAGGCTCTGAGATCTCAGTGGCTCCA 1200
DB 1141 ACCAGAGTGTGCTTCTGCTGCTCTTATGAGAGGCTCTGAGATCTCAGTGGCTCCA 1200
QY 1201 ACAAGTCCCTTGCCTTCATCTGATACAAAACATCTGCAATCCCGCTCAGTATTTGC 1260
DB 1201 ACAAGTCCCTTGCCTTCATCTGATACAAAACATCTGCAATCCCGCTCAGTATTTGC 1260
QY 1261 ACCTTGTGCTAATTAATTAATCTTCCATGACAGCACTCCCTGATCTGAATTCGAGACAGTC 1320
DB 1261 ACCTTGTGCTAATTAATTAATCTTCCATGACAGCACTCCCTGATCTGAATTCGAGACAGTC 1320
QY 1321 TTCTGGAATTTGAGATGATGTTCTTGTGTGCTCTGCACTGATCAAGCTCGATATC 1380
DB 1321 TTCTGGAATTTGAGATGATGTTCTTGTGTGCTCTGCACTGATCAAGCTCGATATC 1380
QY 1381 ACAGAGATGCTGTGACCTGTCTATCTTCTATGAGTTTCGGACCGGCTCAGTCTTTG 1440
DB 1381 ACAGAGATGCTGTGACCTGTCTATCTTCTATGAGTTTCGGACCGGCTCAGTCTTTG 1440
QY 1441 AAGACAGAAACCGGCTTTTGTCAAGCCGACAGCTGATGAATCCGCTTTGTGTG 1500
DB 1441 AAGACAGAAACCGGCTTTTGTCAAGCCGACAGCTGATGAATCCGCTTTGTGTG 1500
QY 1501 GTGTGCTCTTCTGAAAGGGGACATTTGTTATGTTGGAAGAGCCAGAGAGAGAGT 1560
DB 1501 GTGTGCTCTTCTGAAAGGGGACATTTGTTATGTTGGAAGAGCCAGAGAGAGAGT 1560
QY 1561 TACTGAGCCGGAAGATGATGAATTAATCTGGGCTAATCTTTGCTGAAACGGGAATCTAATG 1620
DB 1561 TACTGAGCCGGAAGATGATGAATTAATCTGGGCTAATCTTTGCTGAAACGGGAATCTAATG 1620
QY 1621 GGAAGACCTGTCTTGTGGCAGCTTATATCTGACAGAGATCTCCAGCTGAGACT 1680
DB 1621 GGAAGACCTGTCTTGTGGCAGCTTATATCTGACAGAGATCTCCAGCTGAGACT 1680
QY 1681 TGAACATAGGCTGGAAGAGACTCAAAAGAACCGGCTGAGATTTTGAACAGACACA 1740

QY 356 GGATTAATTGCGAGAGCCACTCTCTAACCCTAATTTGAGCTCCAGAACTCAGAGTGGCT 415
Db 444 GGATTAATTGCGAGAGCCACTCTCTAACCCTAATTTGAGCTCCAGAACTCAGAGTGGCT 503
QY 416 GCTTTTGAATCAACATGCTCAAGGTGCTATTACCCGAAATTCGGAGTGTCAAGAGACTG 475
Db 504 GCTTTTGAATCAACATGCTCAAGGTGCTATTACCCGAAATTCGGAGTGTCAAGAGACTG 563
QY 476 CTTTACCTGAACATCTATGAGGCTGCGCAGCCGATACAGGCTCCAGAGTCCCGCTTT 535
Db 564 CTTTACCTGAACATCTATGAGGCTGCGCAGCCGATACAGGCTCCAGAGTCCCGCTTT 623
QY 536 GGTGTGTTCCCAAGAGTGTCTTCAAGACTGCTCAGCTTCATCTTTGATGAGTCCGC 595
Db 624 GGTGTGTTCCCAAGAGTGTCTTCAAGACTGCTCAGCTTCATCTTTGATGAGTCCGC 683
QY 596 CTTGAGCTGCTTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
Db 684 CTTGAGCTGCTTAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 743
QY 656 TTTTCTTACCAATGAGGATCAGCATGCTCCGAGGAACTGAGCTTCAAGAGCCAGGTGAC 715
Db 744 TTTTCTTACCAATGAGGATCAGCATGCTCCGAGGAACTGAGCTTCAAGAGCCAGGTGAC 803
QY 716 TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
Db 804 TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863
QY 776 CATCTTTGCGAGTCCGCGGAGCCCATAGTGTCTTCTAGCTTATATCTGCTCCCATAGC 835
Db 864 CATCTTTGCGAGTCCGCGGAGCCCATAGTGTCTTCTAGCTTATATCTGCTCCCATAGC 923
QY 836 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTCTTACCTGGA 895
Db 924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTCTTACCTGGA 983
QY 896 GGGCCATGATTTAGAGAGAGTGAAGACCTGAGAGTGTGCACTTTCTGTGTGAACA 955
Db 984 GGGCCATGATTTAGAGAGAGTGAAGACCTGAGAGTGTGCACTTTCTGTGTGAACA 1043
QY 956 TGGCTGAGACTGTGAGGCTGCTGAGGTGCTGAGGACAAACCTCTCAAGAGAGTCT 1015
Db 1044 TGGCTGAGACTGTGAGGCTGCTGAGGTGCTGAGGACAAACCTCTCAAGAGAGTCT 1103
QY 1016 GAGCCCTGCGCAGAAACAAAGTCTTTCACTGAGTGTGATGAGTGTCTTTCTTAA 1075
Db 1104 GAGCCCTGCGCAGAAACAAAGTCTTTCACTGAGTGTGATGAGTGTCTTTCTTAA 1163
QY 1076 TGAAGCTCTAGATCTATTGCTCAGAAAGCATTTAAAGCATTCCTTCATCATCGAGT 1135
Db 1164 TGAAGCTCTAGATCTATTGCTCAGAAAGCATTTAAAGCATTCCTTCATCATCGAGT 1223
QY 1136 CAATTAACAAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1195
Db 1224 CAATTAACAAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1283
QY 1196 CTCCCAAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
Db 1284 CTCCCAAGAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343
QY 1256 TTTGACACTTGTGCTATGATTAATCTTCATGACAGCACTCCCTGACTGAATTCGAGA 1315
Db 1344 TTTGACACTTGTGCTATGATTAATCTTCATGACAGCACTCCCTGACTGAATTCGAGA 1403
QY 1316 CAGCTCTCTGGAATTTGCTGGAAGATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1375
Db 1404 CAGCTCTCTGGAATTTGCTGGAAGATGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1463
QY 1376 ATATCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1435
Db 1464 ATATCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523
QY 1436 CTTTGAAGACAGAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTTGT 1495

Db 1524 CTTTGAAGACAGAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTTGT 1583
QY 1496 GTTCGATGATGCTCTCTGGAAGGGGGAATTTGTTATGTTTCAAGAGGACCAAGAGAGA 1555
Db 1584 GTTCGATGATGCTCTCTGGAAGGGGGAATTTGTTATGTTTCAAGAGGACCAAGAGAGA 1643
QY 1556 GAAATTAATGAGCCGGAAGATGATGAATTAATCTGAGTACCTTTGCTGAACCGGGAATCC 1615
Db 1644 GAAATTAATGAGCCGGAAGATGATGAATTAATCTGAGTACCTTTGCTGAACCGGGAATCC 1703
QY 1616 TAAATGGAACAGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1675
Db 1704 TAAATGGAACAGACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1763
QY 1676 GGAATTAATGAGCCCTGGAAGAGACTCAAGAGCCGAGGAGTGAATTTGAGACAG 1735
Db 1764 GGAATTAATGAGCCCTGGAAGAGACTCAAGAGCCGAGGAGTGAATTTGAGACAG 1823
QY 1736 CACCATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
Db 1824 CACCATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1883
QY 1796 TTTCTCTCTCTCTGAGGCTTTCTTTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1855
Db 1884 TTTCTCTCTCTCTGAGGCTTTCTTTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1943
QY 1856 GTGATTTGCTTCT 1915
Db 1944 GTGATTTGCTTCT 2003
QY 1916 GCTCAGCTGCTTCTCTGAGGAGATCTCTGGAACCAAGCTCTTGGCTGATTTATGAG 1975
Db 2004 GCTCAGCTGCTTCTCTGAGGAGATCTCTGGAACCAAGCTCTTGGCTGATTTATGAG 2063
QY 1976 ACTTGAAGATGATCTTCAAGATCTTTTCAACATCAAAAGTGAATTTGCTTGA 2035
Db 2064 ACTTGAAGATGATCTTCAAGATCTTTTCAACATCAAAAGTGAATTTGCTTGA 2123
QY 2036 GGAACAAAGTTTCTTCAATTAATTTGGAAGAGGCTGCTTATGATTTGATATAT 2095
Db 2124 GGAACAAAGTTTCTTCAATTAATTTGGAAGAGGCTGCTTATGATTTGATATAT 2183
QY 2096 GGTTTGTAATCATATGAATTAATCAAGATGTAATTAAGAAAA 2144
Db 2184 GGTTTGTAATCATATGAATTAATCAAGATGTAATTAAGAAAA 2232

RESULT 5
US-10-047-249
; Sequence 249, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-249

Query Match 88.8%; Score 1917; DB 17; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 5; Indels 10; Gaps 1;
QY 186 GGGCTTCTGCTGAAGGCCACAGAGAACACAGGCTGGAGTGAATTCAGGCCAAGCAA 245

138 GGGCCCTTCTGCTGAAGGGGCAAGAGAAACCAAGGCTGGATGATTCAGGGCAAGCAA 197
139 GTCATGCTGGGGAAGCCCTGCTGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
198 GTGATGCTGGGGAAGCCCTGCTGCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
306 CCCCCGCTGGGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 365
258 CCCCCGCTGGGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 317
366 GAGAGAGCACTCTCTACCTTAATTT-----GTGCTCTCGAAGCTCAGAGTGGCT 415
318 GAGAGAGCACTCTCTACCTTAATTTGTAAGACAGAGTGGCTCCAGAACTCAGAGTGGCT 377
416 GCTCTTAATCAACACATGCTCAAGTGTGATTAACCGGAATTCGGAAGTGTCAAGATCTG 475
378 GCTCTTAATCAACACATGCTCAAGTGTGATTAACCGGAATTCGGAAGTGTCAAGATCTG 437
476 CCTTACCTGAACATCTATGCGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535
438 CCTTACCTGAACATCTATGCGCTGCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 497
536 GGTGTGCTTCCAGAGAGTGGCTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
498 GGTGTGCTTCCAGAGAGTGGCTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557
596 CCTGCTGCTTCAAGAGAGTGGCTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 655
558 CCTGCTGCTTCAAGAGAGTGGCTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
656 TTTCTTCAACATGAGGATCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
618 TTTCTTCAACATGAGGATCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
716 TGTCTGCTCTGCGGCTCAAGAGAGTGGCTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTG 775
678 TGTCTGCTCTGCGGCTCAAGAGAGTGGCTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTG 737
776 CATCTTTGGCGAGTCCGCGGAGCAATGATGTTTCTGATCTTAATCTGCTGCTGCTGCTG 835
738 CATCTTTGGCGAGTCCGCGGAGCAATGATGTTTCTGATCTTAATCTGCTGCTGCTGCTG 797
836 CAAAGGCTTATTCACAAAGGATCATGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895
798 CAAAGGCTTATTCACAAAGGATCATGAGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
896 GGGCCATGATTAAGAGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955
858 GGGCCATGATTAAGAGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917
956 TGTGCTGAGCTTGTAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1015
918 TGTGCTGAGCTTGTAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 977
1016 GACCTTCAAGGAGAAACAAAGTCTTCACTGAGTGTGATGATGATGATGATGATGATGATG 1075
978 GACCTTCAAGGAGAAACAAAGTCTTCACTGAGTGTGATGATGATGATGATGATGATGATG 1037
1076 TGAGGCTTAAGTCTATGCTCAGAAAGATTTAAAGCAATTCCTTCATCATCGAGT 1135
1038 TGAGGCTTAAGTCTATGCTCAGAAAGATTTAAAGCAATTCCTTCATCATCGAGT 1097
1136 CAATACCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1195
1098 CAATACCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157
1136 CTGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1255
1158 CTGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1217
1256 TTTGCACTTGTGTGCTTAATGATCTTCATGCAAGCACTGCTGATGAAATCCGAGA 1315

1218 TTTGCACTTGTGTGCTTAATGATCTTCATGCAAGCACTGCTGATGAAATCCGAGA 1277
1316 CAGCTTCTGAGACTTGTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1375
1278 CAGCTTCTGAGACTTGTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1337
1376 ATATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1435
1338 ATATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1397
1436 CTTTGAACACAGAGCCGCTTGTGTCAAGCCGACCAAGCTGATGATGATGATGATGATGATG 1495
1398 CTTTGAACACAGAGCCGCTTGTGTCAAGCCGACCAAGCTGATGATGATGATGATGATGATG 1457
1496 GTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1555
1458 GTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1517
1556 GAAATTAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1615
1518 GAAATTAAGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1577
1616 TAATGGAACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1675
1578 TAATGGAACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1637
1676 GAACTTGAACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1735
1638 GAACTTGAACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1697
1736 CACCAATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
1698 CACCAATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757
1796 TTTCTCTCTCTCTCTGAGCCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1855
1758 TTTCTCTCTCTCTCTGAGCCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1817
1856 GTGATTTGTTTCT 1915
1818 GTGATTTGTTTCT 1877
1916 GCTCAGCTGCTTCT 1975
1878 GCTCAGCTGCTTCT 1937
1976 ACTTGAATGATCTCTTACAGATTTCTTTCAACATCAAAAGTGCATTTGCTTGGAA 2035
1938 ACTTGAATGATCTCTTACAGATTTCTTTCAACATCAAAAGTGCATTTGCTTGGAA 1997
2036 GGCAACAGATTTCTTCAATTAATTTGGAAGAGGCTGCTTATGATGATGATGATGATGATG 2095
1998 GGCAACAGATTTCTTCAATTAATTTGGAAGAGGCTGCTTATGATGATGATGATGATGATG 2057
2096 GATTTTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2130
2058 GATTTTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2092

RESULT 6
US-10-023-515-3
; Sequence 3, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023, 515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18

OY	996	AAACCCCTCAAGAGAGCTGCTGACCCCTCAGCCAGCAAAA	CAAAAGCTTTTCACTCGAGTGGTT	1055
Db	901	AAACCCCTCAAGAGAGCTGCTGACCCCTCAGCCAGCAAAA	CAAAAGCTTTTCACTCGAGTGGTT	960
OY	1056	GATGGTGTCTTCTTCTTAATGAGCCTCAGATCTAATGTCTCAGAAA	AGCATTTAAAGCA	1115
Db	961	GATGGTGTCTTCTTCTTAATGAGCCTCAGATCTAATGTCTCAGAAA	AGCATTTAAAGCA	1020
OY	1116	ATTCTTCATCATATCGAGATCAATTAACA	CCAGAGTGTGCTTCTGCTGCTATGAAGAG	1175
Db	1021	ATTCTTCATCATATCGAGATCAATTAACA	CCAGAGTGTGCTTCTGCTGCTATGAAGAG	1080
OY	1176	GCTCCTGAGATCCTCAGTGGCTCAAA	CAAGTCCCTTGCCCTCATCTGATTAACAAACATC	1235
Db	1081	GCTCCTGAGATCCTCAGTGGCTCAAA	CAAGTCCCTTGCCCTCATCTGATTAACAAACATC	1140
OY	1236	CTGCACATCCCGCTCAAGTATTGCA	CCTTGTGGCTAATGAATACTTCCATGACAAGCAC	1295
Db	1141	CTGCACATCCCGCTCAAGTATTGCA	CCTTGTGGCTAATGAATACTTCCATGACAAGCAC	1200
OY	1296	TCCCTGACTGAATCCGAGACAGTCTTGTGA	CTTGTGAGATGTGTTCTTTGTGCTC	1355
Db	1201	TCCCTGACTGAATCCGAGACAGTCTTGTGA	CTTGTGAGATGTGTTCTTTGTGCTC	1260
OY	1356	CCTGCACTGATCAAGCTCGATATCA	CAGAGATGCTGTGACACCTGTCTAATCTATGAG	1415
Db	1261	CCTGCACTGATCAAGCTCGATATCA	CAGAGATGCTGTGCTACCTGTCTAATCTATGAG	1320
OY	1416	TTTGGGACCGGGCTCAAGTGTCTTGA	AGACACCAAGCCGGCTTTTGTCAAAAGCCGACAC	1475
Db	1321	TTTGGGACCGGGCTCAAGTGTCTTGA	AGACACCAAGCCGGCTTTTGTCAAAAGCCGACAC	1380
OY	1476	GCTGATGAAGTCCGCTTTGTGTGATG	GTGCTTCTGGAAGGGGACATTTGTATGTTTC	1535
Db	1381	GCTGATGAAGTCCGCTTTGTGTGATG	GTGCTTCTGGAAGGGGACATTTGTATGTTTC	1440
OY	1536	GAAAGAGCCACGAGAGGAGGAGAA	ATTACTGAGCCGGAAGTATGAATACTGGGCTAC	1595
Db	1441	GAAAGAGCCACGAGAGGAGGAGAA	ATTACTGAGCCGGAAGTATGAATACTGGGCTAC	1500
OY	1596	TTTGCTCGAACCAGGAATCTCTAAT	TGGGAACGACCTGTCTCTGAGGCGACGCTTATATATCG	1655
Db	1501	TTTGCTCGAACCAGGAATCTCTAAT	TGGGAACGACCTGTCTCTGAGGCGACGCTTATATATCG	1560
OY	1656	ACTAGACAGTAACTCCACAGCTGGA	CTTGAACATGAGCCTCGAACAGAGACTCAAGAACCG	1715
Db	1561	ACTAGACAGTAACTCCACAGCTGGA	CTTGAACATGAGCCTCGAACAGAGACTCAAGAACCG	1620
OY	1716	CGGGTGAATTTTGGACACAGACAC	ATCCCGCCTGATCCTGTGAGCCCTCGACATGCTCCAC	1775
Db	1621	CGGGTGAATTTTGGACACAGACAC	ATCCCGCCTGATCCTGTGAGCCCTCGACATGCTCCAC	1680
OY	1776	AGTCTCTTTTCTTTAATTCTTCTCT	CTCTCTCCAGCCTTTCTTTTCTTTTGTGCT	1835
Db	1681	AGTCTCTTTTCTTTAATTCTTCTCT	CTCTCTCCAGCCTTTCTTTTCTTTTGTGCT	1740
OY	1836	CCCTGA	1841	
Db	1741	CCCTGA	1746	
RESULT 7				
US-10-674-636-3				
; Sequence 3, Application US/10674636				
; Publication No. US20040086922A1				
; GENERAL INFORMATION:				
; APPLICANT: Curtie, Rory A. J.				
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLTERASE				
; FILE REFERENCE: 10448-122001				
; CURRENT APPLICATION NUMBER: US/10/674 336				

; CURRENT FILING DATE: 2003-09-29
 ; PRIOR APPLICATION NUMBER: US/10/023, 515
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 60/256,369
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: 60/279,508
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1746
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-674-636-3

Query Match 80.9%; Score 1746; DB 18; Length 1746;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 ATGCCACAGGAGCTTACTTCACTGCTTCAACAATGAGCTTTTCTGATTCCTCAGGCC 155
 Db 1 ATGCCACAGGAGCTTACTTCACTGCTTCAACAATGAGCTTTTCTGATTCCTCAGGCC 60
 Qy 156 CTGTTGGGACACAGCAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
 Db 61 CTGTTGGGACACAGCAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
 Qy 216 ACCAGGCTGGATGATTCAGGGCAGAGATCTGCTGCTGGAAGCCCTGCTGCTG 275
 Db 121 ACCAGGCTGGATGATTCAGGGCAGAGATCTGCTGCTGGAAGCCCTGCTGCTG 180
 Qy 276 AACGTTCTCTGAGTCCCTTGTGCTGCTCCCGCTGGGATCCCTGCTGATTTAAGAAC 335
 Db 181 AACGTTCTCTGAGTCCCTTGTGCTGCTCCCGCTGGGATCCCTGCTGATTTAAGAAC 240
 Qy 336 CCGAGCTGATCGCCCTGGGATTAATTGCGAGAGCCACTCTTAATTTTGTGTC 395
 Db 241 CCGAGCTGATCGCCCTGGGATTAATTGCGAGAGCCACTCTTAATTTTGTGTC 300
 Qy 396 CTCCAGAACTCAGATGCTGCTCTTAATCAACAGCTCAAGTGCATTAACCGAA 455
 Db 301 CTCCAGAACTCAGATGCTGCTCTTAATCAACAGCTCAAGTGCATTAACCGAA 360
 Qy 456 TTCCGAGTGAAGAGCTGCTCTACCTGAACATCTAGGCTGGCCGACGCGATACA 515
 Db 361 TTCCGAGTGAAGAGCTGCTCTACCTGAACATCTAGGCTGGCCGACGCGATACA 420
 Qy 516 GGCCTCAAGCTCCCGCTTGTGCTGCTCCAGAGGCTGCTTCAAGCTGCTCAGCC 575
 Db 421 GGCCTCAAGCTCCCGCTTGTGCTGCTCCAGAGGCTGCTTCAAGCTGCTCAGCC 480
 Qy 576 TCCATCTTGAATGGTCCCGCTGCTGCTTAAGAGAGCTGCTGCTGCTGCTG 635
 Db 481 TCCATCTTGAATGGTCCCGCTGCTGCTTAAGAGAGCTGCTGCTGCTGCTGCTG 540
 Qy 636 TACCGGCTAGGAATTTGTTCTTCAACAATGAGTCAAGCTCCGGGGAACCTG 695
 Db 541 TACCGGCTAGGAATTTGTTCTTCAACAATGAGTCAAGCTCCGGGGAACCTG 600
 Qy 696 GCCTTCAAGAGCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
 Db 601 GCCTTCAAGAGCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 Qy 756 GGGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
 Db 661 GGGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 Qy 816 CTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 875
 Db 721 CTTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 Qy 876 GCCATCATCCCTTAAGCTGAGGCTCATGATTAAGAGAGTGAAGACCTGAGGCTGTT 935

Db 781 GCCATCATCCCTTAAGCTGAGGCTCATGATTAAGAGAGTGAAGACCTGAGGCTGTT 840
 Qy 936 GCATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 995
 Db 841 GCATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Qy 996 AAACCTCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
 Db 901 AAACCTCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 Qy 1056 GATGCTGCTTCTTCTTAATGAGCTCTGATCTTAATGCTCAGAAAGCATTTAAGCA 1115
 Db 961 GATGCTGCTTCTTCTTAATGAGCTCTGATCTTAATGCTCAGAAAGCATTTAAGCA 1020
 Qy 1116 ATTCTTCATGATGAGGCTCAATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1175
 Db 1021 ATTCTTCATGATGAGGCTCAATTAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 Qy 1176 GCTCCTGAGATCTCAGTGGCTCCAAAGATCCCTTGGCTTCATCTGATTAACAAATC 1235
 Db 1081 GCTCCTGAGATCTCAGTGGCTCCAAAGATCCCTTGGCTTCATCTGATTAACAAATC 1140
 Qy 1236 CTGACATCCGCTCAGTATTTGCACTTGTGCTAATGAATCTTCCATGACAGAC 1295
 Db 1141 CTGACATCCGCTCAGTATTTGCACTTGTGCTAATGAATCTTCCATGACAGAC 1200
 Qy 1296 TCCCTGATGAAATCCGAGACAGTCTTCTGCACTTGTGCTGAGATGCTTCTTGTGCT 1355
 Db 1201 TCCCTGATGAAATCCGAGACAGTCTTCTGCACTTGTGCTGAGATGCTTCTTGTGCT 1260
 Qy 1356 CCTGACATGATCAAGCTCGATATCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 1415
 Db 1261 CCTGACATGATCAAGCTCGATATCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 Qy 1416 TTTGGGACCGGCTCAGTCTTTGAGAGACAGAGCCGCTTTTGTCAAGCCGACAC 1475
 Db 1321 TTTGGGACCGGCTCAGTCTTTGAGAGACAGAGCCGCTTTTGTCAAGCCGACAC 1380
 Qy 1476 GCTGATGAAGTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1535
 Db 1381 GCTGATGAAGTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
 Qy 1536 GAAAGGACCAAGAGAGAGAGATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1595
 Db 1441 GAAAGGACCAAGAGAGAGAGAGATTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
 Qy 1596 TTTGCTGAACCGGGAATCTTAATGGAAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 1655
 Db 1501 TTTGCTGAACCGGGAATCTTAATGGAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 1560
 Qy 1656 ACTGAGCAGTACCTCAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1715
 Db 1561 ACTGAGCAGTACCTCAGCTGGAATTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1620
 Qy 1716 GGGGAGATTTTGGACAGAGACCATCCCGTGAATCTGCTGCTGCTGCTGCTGCTGCTG 1775
 Db 1621 GGGGAGATTTTGGACAGAGACCATCCCGTGAATCTGCTGCTGCTGCTGCTGCTGCTG 1680
 Qy 1776 AGTCTCTTCTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
 Db 1681 AGTCTCTTCTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
 Qy 1836 CCTTGA 1841
 Db 1741 CCTTGA 1746

RESULT 8
 US-10-451-168-40
 ; Sequence 40, Application US/10451168
 ; Publication No. US20040091969A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SMITHKLINE BEECHAM CORPORATION


```

; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: G950039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 1728
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-451-168-40

```

```

Query Match      76.6%; Score 1652.8; DB 18; Length 1728;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 186 GGGCCCTTCTGCTGAAGGGCCACAGAGAACACACAGGCTGGATTCAGAGGCAAGCAA 245
Db 73 GGGCCCTTCTGCTGAAGGGCCACAGAGAACACACAGGCTGGATTCAGAGGCAAGCAA 132
QY 246 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTTCTCTGAGAGTCCCTTTGCTGCT 305
Db 133 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTTCTCTGAGAGTCCCTTTGCTGCT 192
QY 306 CCCCCGCTGGGATCCCTCGCATTTTACGAACCCGACGCTGCATGGCCCTGGAGATTAATTG 365
Db 193 CCCCCGCTGGGATCCCTCGCATTTTACGAACCCGACGCTGCATGGCCCTGGAGATTAATTG 252
QY 366 CGAAGAGCACCCTCTTACCTTAATTGTGCTCCAGAACTCAGAGTGGCTGCTTTAGAT 425
Db 253 CGAAGAGCACCCTCTTACCTTAATTGTGCTCCAGAACTCAGAGTGGCTGCTTTAGAT 312
QY 426 CAACACATGCTCAAGATGCTATTAACCCGAATTCGAGGTGTCAAGAACTGCTCTTACTG 485
Db 313 CAACATATGCTCAAGATGCTATTAACCCGAATTCGAGGTGTCAAGAACTGCTCTTACTG 372
QY 486 AACATCTATGCGCTGCGCCACAGCCGATACAGGCTCAAGCTCCCCGCTTGTGTGTTC 545
Db 373 AACATCTATGCGCTGCGCCACAGCCGATACAGGCTCAAGCTCCCCGCTTGTGTGTTC 432
QY 546 CCAGAGAGTGTCTTCAAGACTGCTGCTCAGCTTCATCTTTGATGGGTCCGCTGCTGCTG 605
Db 433 CCAGAGAGTGTCTTCAAGACTGCTGCTCAGCTTCATCTTTGATGGGTCCGCTGCTGCTG 492
QY 606 TATGAGAGCCTGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
Db 493 TATGAGAGCCTGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 552
QY 666 ACATGGAGTACAGATGTCTCGGGGAACTGGGCTTCAAGGACCAAGATGCTGTCTGTCTC 725
Db 553 ACATGGAGTACAGATGTCTCGGGGAACTGGGCTTCAAGGACCAAGATGCTGTCTGTCTC 612
QY 726 TGGGTCCAGAAAGAACATCGAGTTCTTGTGGTGGAGCCCAAGCTGTGTGACCATCTTTGGC 785

```

```

Db 613 TGGGTCCAGAAAGAACATCGAGTTCTTGTGGTGGAGCCCAAGCTGTGTGACCATCTTTGGC 672
QY 786 GAGTCCGCGGGAGCCATTAAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 845
Db 673 GAGTCCGCGGGAGCCATTAAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 732
QY 846 TTCCAAGAGCCATGATGAGAGTGGGTGGCCATCATCCCTTACCTGTGAGAGCCCATAT 905
Db 733 TTCCAAGAGCCATGATGAGAGTGGGTGGCCATCATCCCTTACCTGTGAGAGCCCATAT 792
QY 906 TATGAGAAAGTGAAGACCTCGAGGTGTTGCACATTTCTGTGTAAACATGCTGTAGAC 965
Db 793 TATGAGAAAGTGAAGACCTCGAGGTGTTGCACATTTCTGTGTAAACATGCTGTAGAC 852
QY 966 TCTGAGGCGCTGCTGAGGTGCTGAGGCAAAACCCCTCCAAAGAGCTGTGACCTCAGC 1025
Db 853 TCTGAGGCGCTGCTGAGGTGCTGAGGCAAAACCCCTCCAAAGAGCTGTGACCTCAGC 912
QY 1026 CAGAAACAAAGTCTTTCACTGAGGTGATGATGATGATGATGATGATGATGATGATGAT 1085
Db 913 CAGAAACAAAGTCTTTCACTGAGGTGATGATGATGATGATGATGATGATGATGATGAT 972
QY 1086 GATCTATTGTCTGAGAAAGCAATTAAGCAATTCCTTCATCATCGAGTGTCAATTAACAC 1145
Db 973 GATCTATTGTCTGAGAAAGCAATTAAGCAATTCCTTCATCATCGAGTGTCAATTAACAC 1032
QY 1146 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
Db 1033 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
QY 1206 TCCCTTGGCCCTCCATCTGATTAACAAACATCTCTCAATCCCGCTCAGATTTTGACCTT 1265
Db 1093 TCCCTTGGCCCTCCATCTGATTAACAAACATCTCTCAATCCCGCTCAGATTTTGACCTT 1152
QY 1266 GTGGCTAATGAATACCTTCCATGAACAAGCACTCCCTGACGTAAATCCGAGACAGCTT 1325
Db 1153 GTGGCTAATGAATACCTTCCATGAACAAGCACTCCCTGACGTAAATCCGAGACAGCTT 1212
QY 1326 GACTGTGCTGAGATGTGTTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1385
Db 1213 GACTGTGCTGAGATGTGTTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1272
QY 1386 GATGCTGTGACACCTGCTCTAATTAATGATTTTGGCAACCGGCTCAGTGTCTTGAAGC 1445
Db 1273 GATGCTGTGACACCTGCTCTAATTAATGATTTTGGCAACCGGCTCAGTGTCTTGAAGC 1332
QY 1446 AGCAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCGCTTGTGTGCTGTGT 1505
Db 1333 AGCAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCGCTTGTGTGCTGTGTGT 1392
QY 1506 GCTTCTGAAGGGGGGCAATTTTATGTTTGAAGAGCCAGGAGAGAGAGATTAATCTG 1565
Db 1393 GCTTCTGAAGGGGGGCAATTTTATGTTTGAAGAGCCAGGAGAGAGAGATTAATCTG 1452
QY 1566 ASCCGAAGATGATGAATACTGGGCTACTTGTGCTGAAACCGGGAATCTTAATGGGAA 1625
Db 1453 ASCCGAAGATGATGAATACTGGGCTACTTGTGCTGAAACCGGGAATCTTAATGGGAA 1512
QY 1626 GACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1685
Db 1513 GACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1572
QY 1686 ATGAGCCTCGACAGAGACTCAAGAAACCGGCGGGTGAATTTTGAACAGAGACATCCCC 1745
Db 1573 ATGAGCCTCGACAGAGACTCAAGAAACCGGCGGGTGAATTTTGAACAGAGACATCCCC 1692
QY 1746 CTGATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1805
Db 1693 CTGATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692
QY 1806 CTGCTCCAGCCTTCTTTTCTTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1841

```

Db 1693 CTCCTCAGCCTTCTTTTCTTTTGTGCTCTGA 1728

RESULT 9
US-10-094-749-736
; Sequence 736, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 736
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-736

Query Match 73.2%; Score 1579.4; DB 17; Length 1962;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

QY 186 GGGCTTCTGTGAGAGGGCCACAGAGAACACAGAGCTGGATGATTCAGGGCAAGCA 245
Db 222 GGGCTTCTGTGAGAGGGCCACAGAGAACACAGAGCTGGATGATTCAGGGCAAGCA 281
QY 246 GTCACTGTGTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTGTGCTG 305
Db 282 GTCACTGTGTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTGTGCTG 341
QY 306 CCCCCGTGGAGATCCCTGGAATTTAGAACCCGCAAGCCGATGCGCCCTGGAGTAATTTG 365
Db 342 CCCCCGTGGAGATCCCTGGAATTTAGAACCCGCAAGCCGATGCGCCCTGGAGTAATTTG 401
QY 366 CGAGAGCCACTCTTACCTTAATTTGTGCTTCAGAACTCAGATGGCTGTCTTAAGAT 425
Db 402 CGAGAGCCACTCTTACCTTAATTTGTGCTTCAGAACTCAGATGGCTGTCTTAAGAT 461
QY 426 CAACATGCTCAGAGTGTGATTAACCGAAATTTGGAGTTCAGAAAGTGTCTTAACTG 485
Db 462 CAACATGCTCAGAGTGTGATTAACCGAAATTTGGAGTTCAGAAAGTGTCTTAACTG 521
QY 486 AACATCTATGGGCTGCCAGGCGATACAGGCTCAAGCTCCCGTCTTGGTGGTTG 545
Db 522 AACATCTATGGGCTGCCAGGCGATACAGGCTCCAGGCTCCAGGCTCCAGGCTCC 581
QY 546 CCAGAGGTGCTTCAAGACTGTGCTCAGGCTCATCTTGTGATGGTCCGCTGTGCTGC 605
Db 582 CCAGAGGTGCTTCAAGACTGTGCTCAGGCTCATCTTGTGATGGTCCGCTGTGCTGC 641

QY 606 TATGAGAGCTGTGCTGTGTGTGCTGTCAGTACCGGCTAGGAATATTGTGTTCTTAC 665
Db 642 TATGAGAGCTGTGCTGTGTGTGCTGTCAGTACCGGCTAGGAATATTGTGTTCTTAC 701
QY 666 ACATGGGATCAGCATGCTCCGGGGAATCGGGCTTCAAGAGCAGATGGCTGTGCTC 725
Db 702 ACATGGGATCAGCATGCTCCGGGGAATCGGGCTTCAAGAGCAGATGGCTGTGCTC 761
QY 726 TGGGTCCAGAGAAATATGAGTCTTCCGTGGGAGCCCACTGTGTGACATCTTTGGC 785
Db 762 TGGGTCCAGAGAAATATGAGTCTTCCGTGGGAGCCCACTGTGTGACATCTTTGGC 821
QY 786 GAGTCCGGGGAGCATATAGTGTCTTGTCTTATAGTCTTCCCATGGCCAAAGCTTA 845
Db 822 GAGTCCGGGGAGCATATAGTGTCTTGTCTTATAGTCTTCCCATGGCCAAAGCTTA 881
QY 846 TTCCAAGAGCCATGATGAGAGTGGGGTGGCCATCATTCCTTACCTGAGGCCATGAT 905
Db 882 TTCCAAGAGCCATGATGAGAGTGGGGTGGCCATCATTCCTTACCTGAGGCCATGAT 941
QY 906 TATGAGAGAGTGAAGACTGTGAGGTGTGACATTTCTGTGTAAACAATGCTTCAGC 965
Db 942 TATGAGAGAGTGAAGACTGTGAGGTGTGACATTTCTGTGTAAACAATGCTTCAGC 1001
QY 966 TCTGAGGCCCTGTGAGGTGCTGAGGACCAAAACCTCCAGAGAGTGTGACCCCTGAGC 1025
Db 1002 TCTGAGGCCCTGTGAGGTGCTGAGGACCAAAACCTCCAGAGAGTGTGACCCCTGAGC 1061
QY 1026 CAGAAAACAAAGTCTTTGATCTGAGTGTGATGAGTGTGCTTCTTCTTAATGAGCTCTA 1085
Db 1062 CAGAAAACAAAGTCTTTGATCTGAGTGTGATGAGTGTGCTTCTTCTTAATGAGCTCTA 1121
QY 1086 GATCTATTGTCTCAGAAAGCATTTAAGCATTCCTTCATATGAGACTCAATAACAC 1145
Db 1122 GATCTATTGTCTCAGAAAGCATTTAAGCATTCCTTCATATGAGACTCAATAACAC 1181
QY 1146 GAGTGTGCTTCCGTGCTGCTGATGAGAGGCTCTGAGATCTGAGTGGCTCAACAG 1205
Db 1182 GAGTGTGCTTCCGTGCTGCTGATGAGAGGCTCTGAGATCTGAGTGGCTCAACAG 1241
QY 1206 TCCCTTGGCCCTCATCTGATACAAACATCTGACATCCCGCTCAGTATTTGACCTT 1265
Db 1242 TCCCTTGGCCCTCATCTGATACAAACATCTGACATCCCGCTCAGTATTTGACCTT 1301
QY 1266 GTGGCTAATGAATCTTCCATGACAGCACTCCGACTGAATCCGACAGTCTTCTG 1325
Db 1302 GTGGCTAATGAATCTTCCATGACAGCACTCCGACTGAATCCGACAGTCTTCTG 1361
QY 1326 GACTTGTGAGAGTGTCTTGTGCTTCCGTGACCTGATGACAGCTGATATACAGA 1385
Db 1362 GACTTGTGAGAGTGTCTTGTGCTTCCGTGACCTGATGACAGCTGATATACAGA 1421
QY 1386 GATGCTGTGACACTGTCTACTTCTATGAGTTCCGCAACCGGCTCAGTGTGTAAGAC 1445
Db 1422 ----- 1421
QY 1446 ACGAAGCCGCTTTTGTGAAGCCGACCAAGCTGATGAATGTCGCTTGTGTCGTGT 1505
Db 1422 ----- 1421
QY 1506 GCTTCTCGAAGGGGAGCATTTGTTATGTTGGAAGAGCCAGGAGAGAGAGTAACTG 1565
Db 1422 -----GAAAGAGCCAGGAGAGAGAGTAACTG 1451
QY 1566 AGCCGGAAGATGATGAATACTAGGCTACCTTTGTCTGAAACCGGAAATCTTAAGGAAC 1625
Db 1452 AGCCGGAAGATGATGAATACTAGGCTACCTTTGTCTGAAACCGGAAATCTTAAGGAAC 1511
QY 1626 GACTGTCTCTGTGGCCAGCTTATATCTGACTGAGAGTACTTCAAGTGAATTTGAAC 1685
Db 1512 GACTGTCTCTGTGGCCAGCTTATATCTGACTGAGAGTACTTCAAGTGAATTTGAAC 1571

QY 1886 ATGACCTCGGACAGAGACTCAAGAACCGGGGGGATTTTGGACACGACATGCC 1745
 Db 1572 ATGAGCCTCGGACAGAGACTCAAGAACCGGGGGGATTTTGGACACGACATGCC 1631
 QY 1746 CTGATCTGTCTGCTCGGACATGCTCCAGTCTCTTCTTCTTAACTTCTCTCT 1805
 Db 1632 CTGATCTGTCTGCTCGGACATGCTCCAGTCTCTTCTTCTTAACTTCTCTCT 1691
 QY 1806 CTCTCCAGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1865
 Db 1692 CTCTCCAGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1751
 QY 1866 TTTCCTCTTCTCTCCCAATATTTCTCCGCAATAGCTTCTTCTTCTTCTTCTTCT 1925
 Db 1752 TTTCCTCTTCTCTCCCAATATTTCTCCGCAATAGCTTCTTCTTCTTCTTCTTCT 1811
 QY 1926 TTTCCTATGGGGATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1985
 Db 1812 TTTCCTATGGGGATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1871
 QY 1986 GATCCTTACGAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2045
 Db 1872 GATCCTTACGAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1931
 QY 2046 TTTCCTCAATTAATTTGGAAGAGGGCTGGCC 2076
 Db 1932 TTTCCTCAATTAATTTGGAAGAGGGCTGGCC 1962

RESULT 10

US-10-114-270-195
 / Sequence 195, Application US/10114270
 / Publication No. US20040030110A1
 / GENERAL INFORMATION:
 / APPLICANT: Guo, Xiaojia
 / APPLICANT: Kekuda, Ramesh
 / APPLICANT: Miller, Charles E.
 / APPLICANT: Malyskar, Uriel M.
 / APPLICANT: Spylek, Kimberly A.
 / APPLICANT: Patlurajan, Meera
 / APPLICANT: Liu, Ziaohong
 / APPLICANT: Gusev, Vladimir Y.
 / APPLICANT: Li, Li
 / APPLICANT: Vermet, Corine
 / APPLICANT: Zerhusen, Bryan D.
 / APPLICANT: Gorman, Linda
 / APPLICANT: Shenoy, Suresh G.
 / APPLICANT: Pena, Carol E. A.
 / APPLICANT: Smithson, Glenda
 / APPLICANT: Burgess, Catherine E.
 / APPLICANT: Gerlach, Valerie
 / APPLICANT: Padigaru, Muralidhara
 / APPLICANT: Shinkets, Richard A.
 / APPLICANT: Gangolli, Esna A.
 / APPLICANT: Taupier Jr., Raymond J.
 / APPLICANT: Casman, Stacie J.
 / APPLICANT: Ji, Weizhen
 / APPLICANT: Anderson, David W.
 / APPLICANT: Rastelli, Luca
 / APPLICANT: Edinger, Shlomit R.
 / APPLICANT: Stone, David J.
 / APPLICANT: MacDougall, John R.
 / APPLICANT: Rothenberg, Mark E.
 / TITLE OF INVENTION: No. US20040030110A1 Proteins and Nucleic Acids Encoding Same
 / FILING REFERENCE: 21402-322C
 / CURRENT APPLICATION NUMBER: US/10/114, 270
 / PRIOR FILING DATE: 2002-11-27
 / PRIOR APPLICATION NUMBER: 60/281, 086
 / PRIOR FILING DATE: 2001-04-03
 / PRIOR APPLICATION NUMBER: 60/281, 136
 / PRIOR FILING DATE: 2001-04-03
 / PRIOR APPLICATION NUMBER: 60/281, 863

/ PRIOR FILING DATE: 2001-04-05
 / PRIOR APPLICATION NUMBER: 60/281, 906
 / PRIOR FILING DATE: 2001-04-05
 / PRIOR APPLICATION NUMBER: 60/282, 020
 / PRIOR FILING DATE: 2001-04-06
 / PRIOR APPLICATION NUMBER: 60/282, 930
 / PRIOR FILING DATE: 2001-04-10
 / PRIOR APPLICATION NUMBER: 60/282, 934
 / PRIOR FILING DATE: 2001-04-10
 / PRIOR APPLICATION NUMBER: 60/283, 512
 / PRIOR FILING DATE: 2001-04-12
 / PRIOR APPLICATION NUMBER: 60/283, 710
 / PRIOR FILING DATE: 2001-04-13
 / PRIOR APPLICATION NUMBER: 60/284, 234
 / PRIOR FILING DATE: 2001-04-17
 / Remaining Prior Application data removed - See File Wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 470
 / SEQ ID NO: 195
 / LENGTH: 1746
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (1)..(1744)
 / US-10-114-270-195

Query Match 66.5%; Score 1435.6; DB 18; Length 1746;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1496; Conservative 0; Mismatches 59; Indels 15; Gaps 1;

QY 176 GGGAAAACTGGGCTTCTGCTGAAGGCGCAAGAGAACCAAGAGAGAGATGATTC 225
 Db 69 GGTCCAGAGAGGCTTCTGCTGAAGGCGCAAGAGAACCAAGAGAGATGATTC 128
 QY 236 GGGCAAGCAATCACTGTGTGGGAAGCCCTGTGCTGTGAAGTGTCTCGAAGTCC 295
 Db 129 GGGCAAGCAATCACTGTGTGGGAAGCCCTGTGCTGTGAAGTGTCTCGAAGTCC 188
 QY 296 CTTTGTGCTCCCCCGGTGGGAATCCCTGCAATTAAGAACCCGACCTGATGCGCTG 355
 Db 189 CTTTGTGCTCCCCCGGTGGGAATCCCTGCAATTAAGAACCCGACCTGATGCGCTG 248
 QY 356 GGAATTAATTCGAGAACCACTCTTACCTTAATTTGTGCTCAAGATCGAGAGGCT 415
 Db 249 GGAATTAATTCGAGAACCACTCTTACCTTAATTTGTGCTCAAGATCGAGAGGCT 308
 QY 416 GCTTTAGATCAACATGCTCAAGGTGATTAACCGAAATTCGAGGTGCAAGACTG 475
 Db 309 GCTTTAGATCAACATGCTCAAGGTGATTAACCGAAATTCGAGGTGCAAGACTG 368
 QY 476 CCTTACCTGAACATGATGCGCTGCGCAAGCGCATTAAGGCTCAAGCTCCCGTCTT 535
 Db 369 CCTTACCTGAACATGATGCGCTGCGCAAGCGCATTAAGGCTCCCGTCTT 428
 QY 536 GGTGTGTTCCAGAGAGGTGCTTCAAGCTGCTCAAGCTCCATCTTGTAGTGGTCCG 595
 Db 429 GGTGTGTTCCAGAGAGGTGCTTCAAGCTGCTCAAGCTCCATCTTGTAGTGGTCCG 488
 QY 596 CCGGCTGCTTGAAGAGAGTGTGCTGTGCTGCTCAAGTCCGCTTGAAGAAATTTGG 655
 Db 489 CCGGCTGCTTGAAGAGAGTGTGCTGTGCTGCTCAAGTCCGCTTGAAGAAATTTGG 548
 QY 656 TTTCCTCAACATGAGATCAAGATGCTCCGAGAACTGGGCTTCAAGAGACAGATGGC 715
 Db 549 TTTCCTCAACATGAGATCAAGATGCTCCGAGAACTGGGCTTCAAGAGACAGATGGC 608
 QY 716 TGCTGTGCTGTGCTCAAGAGATGAGTCTTGTGTGGGACCCAGCTGTGTAC 775
 Db 609 TGCTGTGCTGTGCTCAAGAGATGAGTCTTGTGTGGGACCCAGCTGTGTAC 668
 QY 776 CATCTTGGCGAGTCCGGGGGAGCCATAAGTGTCTTATCTGTCTCCATGAGC 835
 Db 669 CATCTTGGCGAGTCCGGGGGAGCCATAAGTGTCTTATCTGTCTCCATGAGC 728

Db 613 TGGGTCGAGAAACATCGAGTCTTGGGAGGAGCCCAAGCTGTGACATCTTTGGC 672
Qy 786 GAGTCCGGGAGGACATAGTGTCTTAGTCTTAATCTGTCTCCCAAGGCCATTA 845
Db 673 GAGTCCGGGAGGACATAGTGTCTTAGTCTTAATCTGTCTCCCAAGGCCATTA 732
Qy 846 TTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGGCCCATGAT 905
Db 733 TTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGGCCCATGAT 792
Qy 906 TATGAGAAAGTGAAGACCTGACAGTGGTGGACATTTCTGTGTAACAATGCGTCAGAC 965
Db 793 TATGAGAAAGTGAAGACCTGACAGTGGTGGACATTTCTGTGTAACAATGCGTCAGAC 852
Qy 966 TCTGAGGCGGCGTGAAGTGGTGGAGGACAAACCTTCGAAGAGCTGACCTTCAGC 1025
Db 853 TCTGAGGCGGCGTGAAGTGGTGGAGGACAAACCTTCGAAGAGCTGACCTTCAGC 912
Qy 1026 CAGAAAACAAAGTCTTTCACCTCGAGTGGTGGATGATGCTTCTTCTTAATGAGCCCTTA 1085
Db 913 CAGAAAACAAAGTCTTTCACCTCGAGTGGTGGATGATGCTTCTTCTTAATGAGCCCTTA 972
Qy 1086 GATCTATGCTCTGAGAAAGATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1145
Db 973 GATCTATGCTCTGAGAAAGATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1032
Qy 1146 GAGTGTGGCTTCTGCTGCTATGAAGAGGCTCCGAGATCTCTGAGTGGCTCCAAAG 1205
Db 1033 GAGTGTGGCTTCTGCTGCTATGAAGATTTCTGCTGCTCAATCTGCACTCCCTCA 1092
Qy 1206 TCCCTT-----GCCCTCATCTGAATCAAAACATCTGACATC 1244
Db 1093 AACCGGATGACAGCTTGGCTTCAACAGCTGGGCAATTCACAGAAACATCAGACATC 1152
Qy 1245 CCGCCTCAGTATTTGCACTTGGGCTAATGAATCTTCCATGAACAAGACCTCCGACT 1304
Db 1153 CCGCCTCAGTATTTGCACTTGGGCTAATGAATCTTCCATGAACAAGACCTCCGACT 1212
Qy 1305 GAAATTCGAGACAGTCTTGTGACCTTGTGAGATGTTCTTGTGGTCCCTGACATG 1364
Db 1213 GAAATTCGAGACAGTCTTGTGACCTTGTGAGATGTTCTTGTGGTCCCTGACATG 1272
Qy 1365 ATCAGAGCTCGATATCAAGAGATCTGTGACCTGTCTTATGATGATTTGGGAC 1424
Db 1273 ATCAGAGCTCGATATCAAGAGATCTGTGACCTGTCTTATGATGATTTGGGAC 1332
Qy 1425 CCGCCTCAGTCTTGAAGACAGAGCCGCTTTTGTCAAAAGCCGACCAAGCTGATGA 1484
Db 1333 CCGCCTCAGTCTTGAAGACAGAGCCGCTTTTGTCAAAAGCCGACCAAGCTGATGA 1392
Qy 1485 GTCCGCTTTGTGTGGTGTGCTCTTGAAGGGGACATTTGTATGTTGGAAGAGCC 1544
Db 1393 GTCCGCTTTGTGTGGTGTGCTCTTGAAGGGGACATTTGTATGTTGGAAGAGCC 1452
Qy 1545 ACGGAGAGAGAGTACTAGAGCCGAGAAAGATGAATATCTGGGCTAATCTTGTCTGA 1604
Db 1453 ACGGAGAGAGAGAGTACTAGAGCCGAGAAAGATGAATATCTGGGCTAATCTTGTCTGA 1512
Qy 1605 ACCGGGAATCTTAATGGAAGACCTGTCTGTGGCCAGCTTAATCTGACTGAGCAG 1664
Db 1513 ACCGGGAATCTTAATGGAAGACCTGTCTGTGGCCAGCTTAATCTGACTGAGCAG 1572
Qy 1665 TACCTTCAGTGAAGCTTGAACATGAGCCTCGAACAAGACTCAAAAGAACCGCGAGAGAT 1724
Db 1573 TACCTTCAGTGAAGCTTGAACATGAGCCTCGAACAAGACTCAAAAGAACCGCGAGAGAT 1632
Qy 1725 TTTTGG 1730
Db 1633 GTGTGG 1638

RESULT 12
US-10-381-898-20

Sequence 20, Application US/10381898
Publication No. US20040086887A1
GENERAL INFORMATION:
APPLICANT: AZIMZAI, Yalda, BAUGHN, Mariah R.;
APPLICANT: BOROMSKI, Mark L.; DING, Li;
APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;
APPLICANT: KHAN, Farrah A.; LAU, Preeti G.;
APPLICANT: LEE, Ernestine A.; LU, Dzung Aina M.;
APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
APPLICANT: SANJANMALA, Madhusudan M.;
APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
APPLICANT: CHAMLA, Narinder K.; WALSH, Roderick T.;
APPLICANT: WARREN, Bridget; XU, Yuning;
APPLICANT: YANG, Junning; YAO, Monique; YUE, Henry
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-0233 USN
CURRENT APPLICATION NUMBER: US/10/381, 898
PRIOR FILING DATE: 2003-10-17
PRIOR APPLICATION NUMBER: PCT/US01/30662
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/236, 947
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: US 60/238, 864
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/242, 323
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/247, 581
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/249, 519
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/252, 834
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/250, 567
PRIOR FILING DATE: 2000-11-30
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PERL Program
SEQ ID NO 20
LENGTH: 1857
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7473645CB1
US-10-381-898-20
Query Match 64.2%; Score 1386.2; DB 18; Length 1857;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;
Qy 176 GGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCA 235
Db 264 GGTCCAGAGAGGGCTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCA 323
Qy 236 GGGCAAGCAAGTCACTGTGCTGGAAAGCCCTGTGCTGTGAACCTGTGCTGAGTCCC 295
Db 324 GGGCAAGCAAGTCACTGTGCTGGAAAGCCCTGTGCTGTGAACCTGTGCTGAGTCCC 383
Qy 296 CTTTGTGCTCCCGCTGGAGATCCCTGATTTAGAAACCGCAGCTGATGAGCCCTG 355
Db 384 CTTTGTGCTCCCGCTGGAGATCCCTGATTTAGAAACCGCAGCTGATGAGCCCTG 443
Qy 356 GGAATACTTGGAGAAAGCACTCTTACCTTAATTTGTGCTCCAGAACTGAGTGGCT 415
Db 444 GGAATACTTGGAGAAAGCACTCTTACCTTAATTTGTGCTCCAGAACTGAGTGGCT 503
Qy 416 GCTTTAGATCAACATGCTCAAGTGCATTAACCGAAATTCGAGTGCAGAAAGACTG 475
Db 504 GCTTTAGATCAACATGCTCAAGTGCATTAACCGAAATTCGAGTGCAGAAAGACTG 563
Qy 476 CCTTAACCTGAACATCTATGCGCTGCGCCAGCGGATACAGGCTCAAGCTCCCGTCTT 535

[illegible]

```

Db      1575  GAACTTACTGAGCGCGGAAGAATGATGAATACTAGGGCTACCTTCTCGAACCGGGAAATCC 1634
Ox      1616  TAAATGGAGAGCACTGTCTCTGTGGCCAGCTTAAATCTGACTAGACATACCTTCAGCT 1675
Db      1635  TAAATGGAGAGCACTGTCTCTGTGGCCAGCTTAAATCTGACTAGACATACCTTCAGCT 1694
Ox      1676  GGACTTGAACAATGAGCCTCGGACAGACATCAAGAACCCGGGGGTGATTTTGG 1730
Db      1695  GGACTTGAACAATGAGCCTCGGACAGACATCAAGAACCCGGGGAGAGATGTGG 1749

RESULT 13
US-10-233-933A-3
; Sequence 3, Application US/10233933A
; Publication No. US2004021417A1
; GENERAL INFORMATION:
; APPLICANT: Yamashita, Tetsuro
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
; FILE REFERENCE: SHIG. P202US006
; CURRENT FILING DATE: 2002-09-03
; PRIOR APPLICATION NUMBER: JP2002-057908
; PRIOR FILING DATE: 2002-04-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2145
; TYPE: DNA
; ORGANISM: Felis catus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (175)..(1803)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1804)..(2145)
; OTHER INFORMATION: n is a, c, g, or t
US-10-233-933A-3

Query Match      53.6%; Score 1156; DB 20; Length 2145;
Best Local Similarity 78.2%; Pred. No. 0;
Matches 1466; Conservative 0; Mismatches 366; Indels 22; Gaps 6;

Ox      186  GGGCCTTCTGCTGAAGGGCCACAGAGGAACACAGGCTGGAGTGAATTCAGGGCAAGCA 245
Db      238  GGGCCAGCTGCTGATGACACGATGAGGAGCACAGGCTGGAGTGGATCCGGGGAGAGCA 297
Ox      246  GTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTTCTCGGAAGTCCCTTGTGCT 305
Db      298  ACACTGTACTGGGAAGCAACCGTCCGTGTAACATGTTCTCGGGAATCCCTTAATGCTGA 357
Ox      306  CCCCCGCTGGGATCCCTGCGATTAAAGAACCCGAGCCTGCAATGGCCTGGGATTAATTG 365
Db      358  CTTCTCTAGGGCCCTCTGCGATTTAAGCAACCAAGCTGTCTGCCCCGGAATGACTTC 417
Ox      366  CGAAGAGCACTTCTCACTCACTTAATTGTGCTTCAGAACTCAGAGTGGCTCTTAAAT 425
Db      418  CGAATGCGAATCTCACTCACTTAATTATGCTTCCAGAGCTTAAGATGGCTGTCTCTAT 477
Ox      426  CAACACATGCTCAAGGTGCAATTAACCCGAAATTCGAGATGTCAAGAACTGCTCTACTG 485
Db      478  CAACACGTTCTCAAGGTGCGTTTACCCCAATTTGAAAGGTGCGAAGACTGCTTAACTT 537
Ox      486  AACATCTATGCGCTGCGCCACGACGATTAAGGCTCAAGCTCCCCGATCTGTGTGCTTC 545
Db      538  AACATCTATGCGCGACGACCCATGCGGACAAATGCTCTCAACTCTCTGTCAATGATGTTC 597
Ox      546  CCAGAGGTGCTTCAAGAATGCTCAAGCTCACTCTTAATTTGAATGGGTCCGCTGCTGCC 605
Db      598  CCGGGGGGTGCTTCAAGATGGGTCAAGCTTCTCTTCATGATGGGTCCGCTTGGCTGCC 657
Ox      606  TATAGACGTGCTGTGTGTGTGTGTGTCAAGTACCGGCTAAGAAATTTGGTTTCTTACC 665

```


Db	364	AAcATCTATGGGcCAAGcCCcATGcCGcGAAATGgGCTcCAACCTCCCTGcATGgTgTGCTTC	422
Oy	546	CCAGAGGTCCTCTCAAGACTGgGCTCAGCTTCATCTTTATGGgTCCGCTTGCTgCC	605
Db	424	CCCGGGgGTGCTTCAGATGgGGCTCAGCTTCTCTTCGATGGgGTCCGCTTGCTgCC	483
Oy	606	TATGAGGAcGgTcGTgTgTgGTGCTCAGTACCGGCTAGGAATATTGgTTCCTACc	665
Db	484	TACGAGAcGTCTGATCTGACTACCGACTACCGGCTAGGAATATTGgTTCCTACc	543
Oy	666	ACATGGATcAGACTGCTCCGgGGAcCTGGcCTTCAGAGAcCAGgTGGCTGTCTgCC	725
Db	544	ACAGGGATGAGcATGcCCCGGGGAcCTGGcCTTCGAGAcCAGgTGGCTGTCTgCC	603
Oy	726	TGGGTCcAGAAgAACTGATTCCTCGGTGGGAcCCcAGCTCTGTGAcCATCTTTgGC	785
Db	604	TGGGTCcCGGAcCACTGAGTTCCTCGGTGGGAcCCcAGCTCTGTGAcCATCTTTgGA	663
Oy	786	GAGTCcCGGGGAcCATGATGTTCTATGCTATACGTCTCCATGGcCCAAAGGCTTA	845
Db	664	GAGTCcGGGGAcCATGATGTTCTCAGCTCATCTGTCTCCCATAGcCCAAAGGCTTA	723
Oy	846	TTCCcAAAGcCATCATGAGAGTGGgGTgGCCATCATCCCTTCCTTGAG---GCCCAT	902
Db	724	TTCCcAAAGcCATCATGAGAGTGGgGTgGCCATCTCGCTTCATGATGAGAcCCCT	783
Oy	903	GATTATGAAAGATGAGAcCTGcAGGTgGTTCACATTCCTGTGTAAcCAATGCTCA	962
Db	784	GGTATGAGAGAAgAAgAATTGGAGGTGGCTGTGCTGTATCTGTGTGCTCAATGGCT	843
Oy	963	GACTCTGAGGcCTCTGAGGTGCTTGAGAcCAAAcCCTTCAGAGAcTCTGAcCCTC	1022
Db	844	GACTCTGTGCTCTGCTGAcGTCTTGAGGcCAAAcCCTTCAGAGAcTCTGAcCCTC	903
Oy	1023	AGcCAGAAAcCAAAgTCTTTCACTCGATGGTGTGATGGTGTCTTCTTCTTAATAGcCT	1082
Db	904	AGcAAGAAcTCAcGTTTCATTCcAGTATGATACCTTTCTTCTCGATGAcCT	963
Oy	1083	CTAGATCTATTGTCTCAGAAAGcATTAAAGcATTCCCTTCATCATCGAGTCATTAc	1142
Db	964	GTAGcCTATTGAcTCAAAAGcATTAAATTCAgTCTCTTATCATCGAGTCATTAc	1023
Oy	1143	CACGAGTGTGCTTCGTGTGCTGCTATGAGAGGCTCTTGATTCCTCAGTGGCTCCAc	1202
Db	1024	CACGAGTGTGCTTCGTGTGCT---CACGAGTTCCTGAGTCCTCGGGGGCTCCAAc	1080
Oy	1203	AAgTCCCTGGcCTCCATCTGATCAAAAcATCCGAcCATCCGGCTCAgTATTGGAc	1262
Db	1081	AAgTCTCTGGcCTCTTATGATACAcGTTCTCGAATATTCCAcCCAGATATTGGAc	1140
Oy	1263	CTGTGTGCTAATGAAATCTTCATGAcCAAGAcCTCCCTGAcCTGAAATCCAGAcGTCCT	1322
Db	1141	CTGTGTGCTAATCTTATCTTCAcAAAGAcCTCCCGGTGAAATACGAGATAGTTCCT	1200
Oy	1323	CTGAcCTTGAGATGTGTCTTTGTGTGCTCCCTGAcCTGATCAcAGCTCGATATCAc	1382
Db	1201	CTGAcCTTGAGATGTGTCTTTGTGTGCTCCCTGGGTGTGACAcGCTCGATATCAc	1260
Oy	1383	AGATATGTGTGAcCTGTCTATCTTATAGTTTCGGAcCCGGCTCAgTCTTTGA	1442
Db	1261	AGATATGTGTGAcCTGTCTATCTTATAGTTTCAAcCCGGCTCAgTCTTTAAc	1320
Oy	1443	GACAcGAAAGcCGCTTTGTCAAAAGcCGAcCGCTGATGAAgTCCGCTTGTGTCCGT	1502
Db	1321	GACAcGAGGcCAGCTTCGTGAAGcCGAcCTGATGAAATCCGCTTGTGTGGA	1380
Oy	1503	GGTGTCTTCTGAAGGGGAcCATTTGATTGTTCAAGAcCCAGAGGAcGAAgTTA	1562
Db	1381	GGTGTCTTCTGAAGGGGAcCATTTGATTGTTCAAGAcCCAGAGGAcGAAATTTG	1440
Oy	1563	CTGAGcCGGAAgATGATGAATATCTGGGCTACTTTGTGAcCCGGGAATCTTAATGGG	1622
Db	1441	CTGAGcAGAAgATGATGAATATCTGGGcCAACTTTGTGTGAcCCGGGAcCCCTTAACGGG	1500

```

QY      1683 AACATGAGCCTCGGCACAGACTCAAGAACA CCGGGGTGAATTTTGGACCAGCACCATC   1742
DB      1661 AGGTGTGAGCGTGGSACAAACTGAAGAGCAAGAGGTGAGTTTTGTAATAATTCATT       1620
OY      1743 CCCG 1746
DB      1621 GTCC 1624

RESULT 15
US-10-451-168-42
; Sequence 42, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; PRIOR FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 1071
TYPE: DNA
ORGANISM: Homo sapiens
US-10-451-168-42

Query Match          47.3%; Score 1021.4; DB 18; Length 1071;
Best Local Similarity    99.9%; Pred. No. 3.8e-284;
Matches 1022; Conservative     0; Mismatches    1; Indels    0; Gaps    0;

OY      819 ATACTGCTGCCAATGGCCAAGGCTTATTTCACA AAGCCATCATGTAGAGTAGGGGTGGCC   878
DB      49  ATTCTGTCTCCCATGSCCAAAGGCTTATTTCACA AAGCCATCATGTAGAGTAGGGGTGGCC   108
OY      879 ATCATCCCTTAACCTGAGGCCCATGATTTATGAGAAGTAGGAGAACCTGCAAGTGTGTGCA   938
DB      109 ATCATCCCTTAACCTGAGGCCCATGATTTATGAGAAGTAGGAGAACCTGCAAGTGTGTGCA   168
OY      939 CATTTCTGTGTGAACAATGGGTGAGACTGTGAGGCCCTGTGTGAGGTGSCCTGAGGACAAA        998
DB      169 CAATTCTGTGTGAACAATGGGTGAGACTGTGAGGCCCTGTGTGAGGTGSCCTGAGGACAAA        228
OY      999 CCCTCCAAAGAGGTGTGAGCCCTACGCGCAGAAAAAA AAAGTCTTYCACTCGAGTGTGTGAT         1058
DB      229 CCTCCCAAGAGGTGTGAGCCCTACGCGCAGAAAAAA AAAGTCTTYCACTCGAGTGTGTGAT         288
OY      1059 GGATGCTTTCTTTCTTAATGAGCCTTAAGATCTATGTCTCTGAGAAAGCATTTAAAGCAATT   1118
```


Db 289 |||||
GATGCTTCTTCTTAATGAGCCTCTAGATCTATTGTCTGAGAAAGCATTTAAAGCAATT 348
QY 1119 |||||
CCTTCGATCATCGGAGATCAATTAACACGAGTGTGGCTTCTGCTGCTATGAAGAGGCT 1178
Db 349 |||||
CCTTCGATCATCGGAGATCAATTAACACGAGTGTGGCTTCTGCTGCTATGAAGAGGCT 408
QY 1179 |||||
CCTGAGATCCTCAGTGTGGCTCAACAAATCCCTGCTCCATCGATGATGAAGAAATCCCTG 1238
Db 409 |||||
CCTGAGATCCTCAGTGTGGCTCAACAAATCCCTGCTCCATCGATGATGAAGAAATCCCTG 468
QY 1239 |||||
CAGATCCCGCTCAGTATTTGCACTTGTGGCTAATGAATCTTCCATGACAAAGCACTCC 1298
Db 469 |||||
CAGATCCCGCTCAGTATTTGCACTTGTGGCTAATGAATCTTCCATGACAAAGCACTCC 528
QY 1299 |||||
CTGACTGAATCCGAGACAGTCTTCTGGAATCTTGGAGATGTGTCTTGTGGTCCCT 1358
Db 529 |||||
CTGACTGAATCCGAGACAGTCTTCTGGAATCTTGGAGATGTGTCTTGTGGTCCCT 588
QY 1359 |||||
GCACTGATCAACAGCTCGATATCAACAGATGTGTGCACTGTCTACTTCTATGAGTTT 1418
Db 589 |||||
GCACTGATCAACAGCTCGATATCAACAGATGTGTGCACTGTCTACTTCTATGAGTTT 648
QY 1419 |||||
CGGCAACGGGCTCAGTGTCTTGAAGACAGAAAGCCGGCTTTGTCAAAGCCGACAGCT 1478
Db 649 |||||
CGGCAACGGGCTCAGTGTCTTGAAGACAGAAAGCCGGCTTTGTCAAAGCCGACAGCT 708
QY 1479 |||||
GATGAAGTCGGCTTGTGTGGTGTGCTTCTGGAAGGGGGAATGTGTATGTTCGAA 1538
Db 709 |||||
GATGAAGTCGGCTTGTGTGGTGTGCTTCTGGAAGGGGGAATGTGTATGTTCGAA 768
QY 1539 |||||
GGAGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATGAATACTGGGCTACTTT 1598
Db 769 |||||
GGAGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATGAATACTGGGCTACTTT 828
QY 1599 |||||
GCTCGAACGGGGAATCTTAATGGGAAACGACCTGTCTGTGGCCAGCTTATATCTGACT 1658
Db 829 |||||
GCTCGAACGGGGAATCTTAATGGGAAACGACCTGTCTGTGGCCAGCTTATATCTGACT 888
QY 1659 |||||
GAGCAGTACCTTCAGCTGGAATTGAACATGAGCTCGACAGAGACTCAAGAAACCGCGG 1718
Db 889 |||||
GAGCAGTACCTTCAGCTGGAATTGAACATGAGCTCGACAGAGACTCAAGAAACCGCGG 948
QY 1719 |||||
GTGGAATTTTGAACAGACACCATCCCTGATCTGTGCTCCGACATGTCTCAGT 1778
Db 949 |||||
GTGGAATTTTGAACAGACACCATCCCTGATCTGTGCTCCGACATGTCTCAGT 1008
QY 1779 |||||
CCTCTTCTCTTAATCTTCTCTCTCTCTCTCCAGCCCTTCTTTTGTGTGCTCT 1838
Db 1009 |||||
CCTCTTCTCTTAATCTTCTCTCTCTCTCTCCAGCCCTTCTTTTGTGTGCTCT 1068
QY 1839 TGA 1841
|||
Db 1069 TGA 1071

Search completed: June 15, 2005, 08:31:05
Job time : 1399.55 secs

Db 241 AGCAAGTCATGTGTGGGAAGCCCTGTGCTGTGTAAGTGTTCCTGGAGTCCCTTTTG 300
 Qy 301 CTGCTCCCGCTGGAGATCCCTGGATTATAGAAACCGGCAAGCTTGCAATCCGCTGGAGATA 360
 Db 301 CTGCTCCCGCTGGAGATCCCTGGATTATAGAAACCGGCAAGCTTGCAATCCGCTGGAGATA 360
 Qy 361 ACTTGGAGAAAGCCACTCTACCTTAATTTGTCTCTCCAGAAATCGAGAGGTGTCT 420
 Db 361 ACTTGGAGAAAGCCACTCTACCTTAATTTGTCTCTCCAGAAATCGAGAGGTGTCT 420
 Qy 421 TAGATCAACACATCTCAAGTGTGATTAACCGAAATTCGAGAGTGTCAAGAAATGCTCT 480
 Db 421 TAGATCAACACATCTCAAGTGTGATTAACCGAAATTCGAGAGTGTCAAGAAATGCTCT 480
 Qy 481 ACCTGAACATCTATGTGCTGCCAGCGGATACAGCTTCAGAGTCCCTGCTTGTGT 540
 Db 481 ACCTGAACATCTATGTGCTGCCAGCGGATACAGCTTCAGAGTCCCTGCTTGTGT 540
 Qy 541 GGTTCACGAGAGTGTCTTCAAGCTGTCAAGCTTCATCTTGTGATGGGTCCGCTGTG 600
 Db 541 GGTTCACGAGAGTGTCTTCAAGCTGTCTCAAGCTTCATCTTGTGATGGGTCCGCTGTG 600
 Qy 601 CTGCTCAATGAGAGCTGT 660
 Db 601 CTGCTCAATGAGAGCTGT 660
 Qy 661 TCACCAATGAGATCAGATGTCTCGGGGAACTGGGCTTCAAGGACAGAGTGTGTGTGT 720
 Db 661 TCACCAATGAGATCAGATGTCTCGGGGAACTGGGCTTCAAGGACAGAGTGTGTGTGT 720
 Qy 721 TGTCTGTGGTCCAGAAACATGTGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 Db 721 TGTCTGTGGTCCAGAAACATGTGATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 Qy 781 TTGGCGAGTCCGGGGAGCATTAAGTGTCTAGTCTTATCTGTCTCCAGTGGCAAG 840
 Db 781 TTGGCGAGTCCGGGGAGCATTAAGTGTCTAGTCTTATCTGTCTCCAGTGGCAAG 840
 Qy 841 GCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACTGTAGAGGCC 900
 Db 841 GCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACTGTAGAGGCC 900
 Qy 901 ATGATTAAGAGAGTGAAGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 Db 901 ATGATTAAGAGAGTGAAGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 Qy 961 CAGACTGTGAGGCTCTGTGAGTGTCTGTAGAGCAAAACCCCTCAAGGAGCTGTGAGCC 1020
 Db 961 CAGACTGTGAGGCTCTGTGAGTGTCTGTAGAGCAAAACCCCTCAAGGAGCTGTGAGCC 1020
 Qy 1021 TCAGCGAAGAAACAAAGTCTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 Db 1021 TCAGCGAAGAAACAAAGTCTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 Qy 1081 CTCTAGATCTATGTGTGAGAAAGCATTTAAAGCAATTCCTTCAATCATCGGATCAATA 1140
 Db 1081 CTCTAGATCTATGTGTGAGAAAGCATTTAAAGCAATTCCTTCAATCATCGGATCAATA 1140
 Qy 1141 ACACGAGT 1200
 Db 1141 ACACGAGT 1200
 Qy 1201 ACAAGTCCCTTCCCTCATCTGTATACAAACATCTGTGACATCCGCTCAAGTATTTGC 1260
 Db 1201 ACAAGTCCCTTCCCTCATCTGTATACAAACATCTGTGACATCCGCTCAAGTATTTGC 1260
 Qy 1261 ACCTTGT 1320
 Db 1261 ACCTTGT 1320
 Qy 1321 TTCTGGAATGCTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
 Db 1321 TTCTGGAATGCTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380

Qy 1381 ACAGAGATCTGTGTGACCTGTCTATCTTCTATGAGTTTGGGACGGGCTGAGTCTTG 1440
 Db 1381 ACAGAGATCTGTGTGACCTGTCTATCTTCTATGAGTTTGGGACGGGCTGAGTCTTG 1440
 Qy 1441 AAGACAGAAAGCTGT 1500
 Db 1441 AAGACAGAAAGCTGT 1500
 Qy 1501 GTGTGTCTTCTGAAAGGGGCAATTTATGTTGTGAAGAGAGCCAGAGAGAGAGAGT 1560
 Db 1501 GTGTGTCTTCTGAAAGGGGCAATTTATGTTGTGAAGAGAGCCAGAGAGAGAGAGT 1560
 Qy 1561 TACTGAGCCGGAAGATGATGAAATCTGAGTACTTGTGTGTGTGTGTGTGTGTGTGT 1620
 Db 1561 TACTGAGCCGGAAGATGATGAAATCTGAGTACTTGTGTGTGTGTGTGTGTGTGTGT 1620
 Qy 1621 GGAAGACCTGT 1680
 Db 1621 GGAAGACCTGT 1680
 Qy 1681 TGAACATGAGCTGTGACAGAGACTCAAAAGAACGGGGTGGATTTTGGACAGAGCA 1740
 Db 1681 TGAACATGAGCTGTGACAGAGACTCAAAAGAACGGGGTGGATTTTGGACAGAGCA 1740
 Qy 1741 TCCCTGT 1800
 Db 1741 TCCCTGT 1800
 Qy 1801 TCTCTCTCTCTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 Db 1801 TCTCTCTCTCTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1860
 Qy 1861 TTGGTTCCT 1920
 Db 1861 TTGGTTCCT 1920
 Qy 1921 GCTGCTTCTTATGAGGATCTTGTGCAAAACAGTCTGTGTGTGTGTGTGTGTGTGT 1980
 Db 1921 GCTGCTTCTTATGAGGATCTTGTGCAAAACAGTCTGTGTGTGTGTGTGTGTGTGT 1980
 Qy 1981 GGAATGATCTTACAGATCTTCTTCAACATCAAAAGGCAATTTGTGTGTGTGTGT 2040
 Db 1981 GGAATGATCTTACAGATCTTCTTCAACATCAAAAGGCAATTTGTGTGTGTGTGT 2040
 Qy 2041 CAAGATTTCTTCAATTAATTTGGAAGAGGCTGTGTGTGTGTGTGTGTGTGTGT 2100
 Db 2041 CAAGATTTCTTCAATTAATTTGGAAGAGGCTGTGTGTGTGTGTGTGTGTGTGT 2100
 Qy 2101 TGTACTCATGAAATTAATTCAGAAATGTAATTAATTAATTAATTAATTAATTAAT 2158
 Db 2101 TGTACTCATGAAATTAATTCAGAAATGTAATTAATTAATTAATTAATTAATTAAT 2158
 RESULT 2
 AX600124 2158 bp DNA linear PAT 14-FEB-2003
 LOCUS Sequence 1 from Patent WO0250256.
 DEFINITION AX600124
 VERSION AX600124
 KEYWORDS AX600124.1 GI:284400190
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 Curtis, R.A. and Silos-Santiago, I.
 53010, a novel human carboxylesterase family member and uses
 thereof
 Patent: WO 0250256-A 1 27-JUN-2002;
 Millennium Pharmaceuticals, Inc. (US)
 JOURNAL
 FEATURES
 source 1..2158

CDS

/organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 96..1841
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD66643.1"
 /db_xref="GI:28400191"
 /translation="MPOGLTSSASQCEFLILQPLIGHROMKGTGSAEGPORNTRLG
 WIOQVAVLQSPVPVNFVGPAPAPLGSRTNPDQSPMDLREASTYDNLQ
 NSEMLDQHLKTHYPRKVSSECLNLTAPAHADGSKLPVLPFGAFKTS
 SIFDSSALAAVEDVAVVQYRLGIFFETIWDQHPGWAFKDOVALSWQKILF
 FGDPSSVTIIFESGAIISVSSILIPMAKGLFKAIMSGVAIIPYLEAHYKESD
 LQVAFPCGNNSDSEALRCRTKPSKELTISOKTSGFTVVDGAFEPNPLDLS
 OKAFKAIPIIIGVNHGEGFLPMKEAPELISGNSKSLALHILIONIHLPGYLHLVA
 NEYFKHSITETIRSDLLDLSDVFPVVPALITARYHNDAGPYFYERHRPCQED
 TKPAPVADHAEVRFVREGAFKQDVIWFEGATEERKLSKMKMTWTFRTGSPN
 GNDLSLWPAYNLTQYQLDLNMSLGRLKEPRVDFWTSTPLILSASDMLSPSL
 TFLSLQPFPPFCAP"

ORIGIN

Query Match 100.0%; Score 2158; DB 6; Length 2158;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CCAGCGCTCCGAAAAACAGCGCTGGAGCAATGTGAGTAAAGCATGTATATAAAAAGA 60
DB 1 CCAGCGCTCCGAAAAACAGCGCTGGAGCAATGTGAGTAAAGCATGTATATAAAAAGA 60
OY 61 TTTAAAAATTATCTTAAATAAAGTACGAATCCCATGCGACAGGAGCTTACTTCATCTG 120
DB 61 TTTAAAAATTATCTTAAATAAAGTACGAATCCCATGCGACAGGAGCTTACTTCATCTG 120
OY 121 CTTCACATGTGCTTTTCTTCTGATTTCTCAGCCCTGTGGGACACAGACAGTGGGGAA 180
DB 121 CTTCACATGTGCTTTTCTTCTGATTTCTCAGCCCTGTGGGACACAGACAGTGGGGAA 180
OY 121 CTTCACATGTGCTTTTCTTCTGATTTCTCAGCCCTGTGGGACACAGACAGTGGGGAA 180
DB 121 CTTCACATGTGCTTTTCTTCTGATTTCTCAGCCCTGTGGGACACAGACAGTGGGGAA 180
OY 181 AAACGTGGGCTTCTGTGAAAGGGCCACAGAGAACAGAGCTGGGATTCAGGGCA 240
DB 181 AAACGTGGGCTTCTGTGAAAGGGCCACAGAGAACAGAGCTGGGATTCAGGGCA 240
OY 181 AAACGTGGGCTTCTGTGAAAGGGCCACAGAGAACAGAGCTGGGATTCAGGGCA 240
DB 181 AAACGTGGGCTTCTGTGAAAGGGCCACAGAGAACAGAGCTGGGATTCAGGGCA 240
OY 241 AGCAAGTCACTGTGCTGGAGAACCCCTGTGCTGTGAAACGTGTCTCGAGAGTCCCTTTG 300
DB 241 AGCAAGTCACTGTGCTGGAGAACCCCTGTGCTGTGAAACGTGTCTCGAGAGTCCCTTTG 300
OY 241 AGCAAGTCACTGTGCTGGAGAACCCCTGTGCTGTGAAACGTGTCTCGAGAGTCCCTTTG 300
DB 241 AGCAAGTCACTGTGCTGGAGAACCCCTGTGCTGTGAAACGTGTCTCGAGAGTCCCTTTG 300
OY 301 CTGCTCCCCCGCTGGGATCCCTGCGATTTCAGAACCCGCGAGCCTGCATGCCCTGGGATA 360
DB 301 CTGCTCCCCCGCTGGGATCCCTGCGATTTCAGAACCCGCGAGCCTGCATGCCCTGGGATA 360
OY 301 CTGCTCCCCCGCTGGGATCCCTGCGATTTCAGAACCCGCGAGCCTGCATGCCCTGGGATA 360
DB 301 CTGCTCCCCCGCTGGGATCCCTGCGATTTCAGAACCCGCGAGCCTGCATGCCCTGGGATA 360
OY 361 ACTTGCGGAGAACCCCTCTACCCCTAATTGTGCTTCAGAACTCAGAGTGGCTGCT 420
DB 361 ACTTGCGGAGAACCCCTCTACCCCTAATTGTGCTTCAGAACTCAGAGTGGCTGCT 420
OY 361 ACTTGCGGAGAACCCCTCTACCCCTAATTGTGCTTCAGAACTCAGAGTGGCTGCT 420
DB 361 ACTTGCGGAGAACCCCTCTACCCCTAATTGTGCTTCAGAACTCAGAGTGGCTGCT 420
OY 421 TAGATCAACACATGTCTCAGGTGCAATTACCCGAAATTGCGAGTGTCAAGAGACTGCTCT 480
DB 421 TAGATCAACACATGTCTCAGGTGCAATTACCCGAAATTGCGAGTGTCAAGAGACTGCTCT 480
OY 421 TAGATCAACACATGTCTCAGGTGCAATTACCCGAAATTGCGAGTGTCAAGAGACTGCTCT 480
DB 421 TAGATCAACACATGTCTCAGGTGCAATTACCCGAAATTGCGAGTGTCAAGAGACTGCTCT 480
OY 481 ACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCAGAGTCCCGCTTGTGTGT 540
DB 481 ACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCAGAGTCCCGCTTGTGTGT 540
OY 481 ACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCAGAGTCCCGCTTGTGTGT 540
DB 481 ACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTTCAGAGTCCCGCTTGTGTGT 540
OY 541 GGTTCGCCAGGAGTGTCTCAAGACTGAGCTCAGCCCTCACTTTGATGGGCTCGCCCTG 600
DB 541 GGTTCGCCAGGAGTGTCTCAAGACTGAGCTCAGCCCTCACTTTGATGGGCTCGCCCTG 600
OY 541 GGTTCGCCAGGAGTGTCTCAAGACTGAGCTCAGCCCTCACTTTGATGGGCTCGCCCTG 600
DB 541 GGTTCGCCAGGAGTGTCTCAAGACTGAGCTCAGCCCTCACTTTGATGGGCTCGCCCTG 600
OY 601 CTGCGCTATGAGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 601 CTGCGCTATGAGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
OY 601 CTGCGCTATGAGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
DB 601 CTGCGCTATGAGGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
OY 661 TCACCAATGGGATCAGATGCTCCGCGGGAATGGGCTTCAGAGACAGAGTGGCTGCTC 720
DB 661 TCACCAATGGGATCAGATGCTCCGCGGGAATGGGCTTCAGAGACAGAGTGGCTGCTC 720

```

```

OY 721 TGTCTGGGTCCAGAGAACATGAGTCTTGTGGTGGGACCCCACTCTGTGACCATCT 780
DB 721 TGTCTGGGTCCAGAGAACATGAGTCTTGTGGTGGGACCCCACTCTGTGACCATCT 780
OY 781 TTGGCGAGTCCCGGGAGGCGATTAAGTGTCTTAACTTAACTGTCTCCCAATGCGCAAG 840
DB 781 TTGGCGAGTCCCGGGAGGCGATTAAGTGTCTTAACTTAACTGTCTCCCAATGCGCAAG 840
OY 841 GCTTATTTCCACAAAGCATCATGAGAGTGGGAGCCATCATCCCTTACTGAGAGCCC 900
DB 841 GCTTATTTCCACAAAGCATCATGAGAGTGGGAGCCATCATCCCTTACTGAGAGCCC 900
OY 841 GCTTATTTCCACAAAGCATCATGAGAGTGGGAGCCATCATCCCTTACTGAGAGCCC 900
DB 841 GCTTATTTCCACAAAGCATCATGAGAGTGGGAGCCATCATCCCTTACTGAGAGCCC 900
OY 901 ATGATTAAGAGAGTGAAGACCTGACAGTGTTCACATTTCTGTGTGTAACATGCGT 960
DB 901 ATGATTAAGAGAGTGAAGACCTGACAGTGTTCACATTTCTGTGTGTAACATGCGT 960
OY 901 ATGATTAAGAGAGTGAAGACCTGACAGTGTTCACATTTCTGTGTGTAACATGCGT 960
DB 901 ATGATTAAGAGAGTGAAGACCTGACAGTGTTCACATTTCTGTGTGTAACATGCGT 960
OY 961 CAGACTGTAGGCTCTGTGAGGTGCTGAGGACAAACCTTCAGAGAGCTGTGACCC 1020
DB 961 CAGACTGTAGGCTCTGTGAGGTGCTGAGGACAAACCTTCAGAGAGCTGTGACCC 1020
OY 1021 TCAGCCAGAAAAAAGTCTTCACTGAGTGGTGAATGATCTTCTTCTTAAATGAGC 1080
DB 1021 TCAGCCAGAAAAAAGTCTTCACTGAGTGGTGAATGATCTTCTTCTTAAATGAGC 1080
OY 1081 CTCTAGATCTATTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATA 1140
DB 1081 CTCTAGATCTATTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATA 1140
OY 1141 ACCAGAGTGTGGCTTCTGTGCTGCTTGAAGAGAGGCTCTGAGATCTCAAGTGTCTCA 1200
DB 1141 ACCAGAGTGTGGCTTCTGTGCTGCTTGAAGAGAGGCTCTGAGATCTCAAGTGTCTCA 1200
OY 1201 ACAAAGTGTGGCTTCTGTGCTGCTTGAAGAGAGGCTCTGAGATCTCAAGTGTCTCA 1200
DB 1201 ACAAAGTGTGGCTTCTGTGCTGCTTGAAGAGAGGCTCTGAGATCTCAAGTGTCTCA 1200
OY 1201 ACAAAGTGTGGCTTCTGTGCTGCTTGAAGAGAGGCTCTGAGATCTCAAGTGTCTCA 1200
DB 1201 ACAAAGTGTGGCTTCTGTGCTGCTTGAAGAGAGGCTCTGAGATCTCAAGTGTCTCA 1200
OY 1261 ACCTGTGCTTAATGAATCTTCATGACAGACACTCCCTGACTGAATAATCCGAGACGT 1320
DB 1261 ACCTGTGCTTAATGAATCTTCATGACAGACACTCCCTGACTGAATAATCCGAGACGT 1320
OY 1261 ACCTGTGCTTAATGAATCTTCATGACAGACACTCCCTGACTGAATAATCCGAGACGT 1320
DB 1261 ACCTGTGCTTAATGAATCTTCATGACAGACACTCCCTGACTGAATAATCCGAGACGT 1320
OY 1321 TTCTGAGCTTGTGAGATGTGTTCTTTGTGTGCTCTGCACTGATCACAGCTCGATATC 1380
DB 1321 TTCTGAGCTTGTGAGATGTGTTCTTTGTGTGCTCTGCACTGATCACAGCTCGATATC 1380
OY 1381 AAGAGAGCTGTGAGACCTGTCTTCTAATGAGTTTCGAGACCGGCTCAGTGTGTTG 1440
DB 1381 AAGAGAGCTGTGAGACCTGTCTTCTAATGAGTTTCGAGACCGGCTCAGTGTGTTG 1440
OY 1381 AAGAGAGCTGTGAGACCTGTCTTCTAATGAGTTTCGAGACCGGCTCAGTGTGTTG 1440
DB 1381 AAGAGAGCTGTGAGACCTGTCTTCTAATGAGTTTCGAGACCGGCTCAGTGTGTTG 1440
OY 1441 AAGAGAGCTGTGAGACCTGTCTTCTAATGAGTTTCGAGACCGGCTCAGTGTGTTG 1500
DB 1441 AAGAGAGCTGTGAGACCTGTCTTCTAATGAGTTTCGAGACCGGCTCAGTGTGTTG 1500
OY 1501 GTGTGCTTCTTGAAGGGGACATTTGTTATGTTGAAAGAGCCACGAGAGAGAGACT 1560
DB 1501 GTGTGCTTCTTGAAGGGGACATTTGTTATGTTGAAAGAGCCACGAGAGAGAGACT 1560
OY 1501 GTGTGCTTCTTGAAGGGGACATTTGTTATGTTGAAAGAGCCACGAGAGAGAGACT 1560
DB 1501 GTGTGCTTCTTGAAGGGGACATTTGTTATGTTGAAAGAGCCACGAGAGAGAGACT 1560
OY 1561 TACTGAGCGGAAAGATGATGAATACTGGGCTTACCTTTGCTGGAACCGGGAATCCTAATG 1620
DB 1561 TACTGAGCGGAAAGATGATGAATACTGGGCTTACCTTTGCTGGAACCGGGAATCCTAATG 1620
OY 1561 TACTGAGCGGAAAGATGATGAATACTGGGCTTACCTTTGCTGGAACCGGGAATCCTAATG 1620
DB 1561 TACTGAGCGGAAAGATGATGAATACTGGGCTTACCTTTGCTGGAACCGGGAATCCTAATG 1620
OY 1621 GGAAGCACTGTCTGTGAGCAGCTTATATCTGACAGAGTACCTTCAGCTGAGACT 1680
DB 1621 GGAAGCACTGTCTGTGAGCAGCTTATATCTGACAGAGTACCTTCAGCTGAGACT 1680
OY 1681 TGAACATGAGCTTCGAGACAGAGACTCAAGAAACCCGAGGTGATTTTGAACAGACCA 1740
DB 1681 TGAACATGAGCTTCGAGACAGAGACTCAAGAAACCCGAGGTGATTTTGAACAGACCA 1740
OY 1741 TCCCTCTGATCTGTCTGCTCCGACATGCTCAAGTCTCTTCTTCTTAACTTTCC 1800
DB 1741 TCCCTCTGATCTGTCTGCTCCGACATGCTCAAGTCTCTTCTTCTTAACTTTCC 1800
OY 1801 TCTCTCTCTCCAGCTTTTCTTTTGTGTGCTCTTGAAGAGTATCTTTCTGTGAT 1860

```

```
Db 1801 TCTGCTCTCCAGCCTTCTTTCTTTGTGCTCTTGAAGATATCTTCTGTGAT 1860
Qy 1861 TTTGGTTTCCTTCTCTCTCCCAATATTTCTCCCAATCATTAAGCTTCTTGAAGCTCA 1920
Db 1861 TTTGGTTTCCTTCTCTCTCCCAATATTTCTCCCAATCATTAAGCTTCTTCTGAAGCTCA 1920
Qy 1921 GCTGCTTCTTAAGGGAATCTTGGCAAAACAAGCTGCTTGGCTATATTTTATGACTTA 1980
Db 1921 GCTGCTTCTTAAGGGAATCTTGGCAAAACAAGCTGCTTGGCTATATTTTATGACTTA 1980
Qy 1981 GGAATGATCTTAAAGAAATCTTTCAACATCAAAAAGTCAATTTCTTGGAGGCA 2040
Db 1981 GGAATGATCTTAAAGAAATCTTTCAACATCAAAAAGTCAATTTCTTGGAGGCA 2040
Qy 2041 CAAGATTTCTTCAATTAATTTGGAAGAGGCTGGCTTATTAAGTTGCTATATATGTTT 2100
Db 2041 CAAGATTTCTTCAATTAATTTGGAAGAGGCTGGCTTATTAAGTTGCTATATATGTTT 2100
Qy 2101 TGTAACTCATATGAATTAATCAGATGTAAATAGAAAAAATTTTAAAAA 2158
Db 2101 TGTAACTCATATGAATTAATCAGATGTAAATAGAAAAAATTTTAAAAA 2158
RESULT 3
AX477711 2232 bp DNA linear PAT 12-AUG-2002
LOCUS AX477711
DEFINITION Sequence 23 from Patent WO0246426.
ACCESSION AX477711
VERSION AX477711.1 GI:22216866
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Sanjanwala, M. M., Yao, M. G., Au-Yang, J., Baughn, M. R., Arvizu, C.,
Rising, H. Z., Lee, E. A., Ding, L., Hafalia, A. J., Tang, Y. T., Yue, H.,
Tribouley, C. M., Lu, D. A., Lal, P. G., Warren, B. A., Yang, J.,
Wallis, N. K., Nguyen, D. B., Gandhi, A. R. and Ison, C. H.
Drug metabolizing enzymes
Patent: WO 0246426-A 23 13-JUN-2002;
Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1..2232
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="caxon:9606"
/note="Incyte ID No: 6538080CB1"
ORIGIN
Query_Match 90.5%; Score 1953; DB 6; Length 2232;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1959; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 176 GGGAAAAAAGCTGGGCTTCTGCTGAAGGGCAAGAGAAACAAGAGCTGGAGATGATTTCA 235
Db 264 GGTCCAAAGAGAGGCTTCTGCTGAAGGGCAAGAGAAACAAGAGCTGGAGATGATTTCA 323
Qy 236 GGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAACGTATTCCTCGAGTCCC 295
Db 324 GGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTGAACGTATTCCTCGAGTCCC 383
Qy 296 CTTTGCTGCTGCTCCCGGCTGGGATCCCGGCAATTAACGAACCGGAGCTGTGATGCGCTTG 355
Db 384 CTTTGCTGCTGCTCCCGGCTGGGATCCCGGCAATTAACGAACCGGAGCTGTGATGCGCTTG 443
Qy 356 GGAATTAAGTGAAGAGCAAGCTTCAAGCTTATTTGAGCTCCAGAACTCAAGATGGCT 415
Db 444 GGAATTAAGTGAAGAGCAAGCTTCAAGCTTATTTGAGCTCCAGAACTCAAGATGGCT 503
Qy 416 GCTTGTGATCAACATGCTCAAGGTGCAATTAACCGAAATTTGGAGTGTCAAGAGCTG 475
Db 504 GCTTGTGATCAACATGCTCAAGGTGCAATTAACCGAAATTTGGAGTGTCAAGAGCTG 563
Qy 476 CCGTAACTGAACATCTAATGAGCGCTGGCCAGCCGATACAGAGCTCAAGCTCCCGTCTT 535
Db 564 CCGTAACTGAACATCTAATGAGCGCGCTGGCCAGCCGATACAGAGCTCAAGCTCCCGTCTT 623
Qy 536 GGTGTGTGTTCCAGAGAGGTGCTTCAAGACTGGCTCAAGCTTCAATTTGATGGGTCCG 595
Db 624 GGTGTGTGTTCCAGAGAGGTGCTTCAAGACTGGCTCAAGCTTCAATTTGATGGGTCCG 683
Qy 596 CCGGTGTGCTTAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 655
Db 684 CCGGTGTGCTTAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 743
Qy 656 TTTCTTCAACATAGGATTCAGATGCTCCGAGGAACTGGGCTTCAAGAGCCAGGTGGC 715
Db 744 TTTCTTCAACATAGGATTCAGATGCTCCGAGGAACTGGGCTTCAAGAGCCAGGTGGC 803
Qy 716 TGCTGTGCTGGGTCCAGAAACAATGAGTCTTGGTGGGAGCCCAAGCTCTGTGAC 775
Db 804 TGCTGTGCTGGGTCCAGAAACAATGAGTCTTGGTGGGAGCCCAAGCTCTGTGAC 863
Qy 776 CATCTTGTGGAGTCCCGGAGGAGCCATTAAGTCTTAACTTATCTGTCTCCATGGC 835
Db 864 CATCTTGTGGAGTCCCGGAGGAGCCATTAAGTCTTAACTTATCTGTCTCCATGGC 923
Qy 836 CAAAGGCTTATTCAGAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTCACTGGA 895
Db 924 CAAAGGCTTATTCAGAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTCACTGGA 983
Qy 896 GGCCCATGATTAAGAGAGTGAAGCTGAGAGCTGAGAGTGTGCAATTTCTGTGTAA 955
Db 984 GGCCCATGATTAAGAGAGTGAAGCTGAGAGTGTGCAATTTCTGTGTAA 1043
Qy 956 TGCTGTGACTGTGAGGCTGCTGAGTGTGCTGAGGACAAACCCCTCAAGAGGCTGCT 1015
Db 1044 TGCTGTGACTGTGAGGCTGCTGAGTGTGCTGAGGACAAACCCCTCAAGAGGCTGCT 1103
Qy 1016 GACCTTGAAGCAAGAAACAAAGTCTTCACTGAGTGTGTGATGTGCTTCTTCTTAA 1075
Db 1104 GACCTTGAAGCAAGAAACAAAGTCTTCACTGAGTGTGTGATGTGCTTCTTCTTAA 1163
Qy 1076 TGAAGCTGTGATCTATGTGCTCAGAAAGATTTAAAGCAATTCCTCATATCGGAGT 1135
Db 1164 TGAAGCTGTGATCTATGTGCTCAGAAAGATTTAAAGCAATTCCTCATATCGGAGT 1223
Qy 1136 CAATTAACCAAGATGTGGCTTCTGTGCTGTGAAGAGAGCTCTGAGATCTCTAGTGG 1195
Db 1224 CAATTAACCAAGATGTGGCTTCTGTGCTGTGAAGAGAGCTCTGAGATCTCTAGTGG 1283
Qy 1196 CTCCAACAAGTCCCTGCTCCATCTGATTAACAAACATCTGACATTCGCGCTGAGTA 1255
Db 1284 CTCCAACAAGTCCCTGCTCCATCTGATTAACAAACATCTGACATTCGCGCTGAGTA 1343
Qy 1256 TTTGCACTTGTGCTAATGAATATCTTCATGACAAGCACTCCCTGACCTGAATCCGAGA 1315
Db 1344 TTTGCACTTGTGCTAATGAATATCTTCATGACAAGCACTCCCTGACCTGAATCCGAGA 1403
Qy 1316 CAGTCTTCTGCACTTGTGAGATGTGTCTTTGTGTGCTCCTGACATGATCAAGCTCG 1375
Db 1404 CAGTCTTCTGCACTTGTGAGATGTGTCTTTGTGTGCTCCTGACATGATCAAGCTCG 1463
Qy 1376 ATATCAAGAGATGTGCTGCACTGTGTAATTAAGTTTGGGACCGGCTCAAGT 1435
Db 1464 ATATCAAGAGATGTGCTGCACTGTGTAATTAAGTTTGGGACCGGCTCAAGT 1523
Qy 1436 CTTTGAAGCAAGAGCGGCTTTTGTAAAGCCGACACGCTGATGAAGTCCGCTTGT 1495
Db 1524 CTTTGAAGCAAGAGCGGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1583
Qy 1496 GTTGTGTGCTTCTGAAAGGGGAGCATTTGTTGTAAGAGGCAAGAGAGGA 1555
Db 1584 GTTGTGTGCTTCTGAAAGGGGAGCATTTGTTGTAAGAGGCAAGAGAGGA 1643
```

QY 1556 GAATTAAGAGCCGGAAGATGATGAAATATGAGGCTACCTTGTCTGAACCGGGAATCC 1615
 DB 1644 GAAAGTACTGAGCCGGAAGATGATGAAATATGAGGCTACCTTGTCTGAACCGGGAATCC 1703
 QY 1616 TAATGGGAACGAACCTGTCTGTGGCAGCTTAAATCTGACTGAGCACTACCTCCAGCT 1675
 DB 1704 TAATGGGAACGAACCTGTCTGTGGCAGCTTAAATCTGACTGAGCACTACCTCCAGCT 1763
 QY 1676 GGAATTGAACATGAGCTCTGGACAGAGACTCAAGAAACCGCGGGTGGATTTTGGACCAAG 1735
 DB 1764 GGAATTGAACATGAGCTCTGGACAGAGACTCAAGAAACCGCGGGTGGATTTTGGACCAAG 1823
 QY 1736 CACCAATCCCCCGTATCCCTGTGCGCCCTGGACATGCGCAAGTCCCTTCTTCCCTTAAC 1795
 DB 1824 CACCAATCCCCCGTATCCCTGTGCGCCCTGGACATGCGCAAGTCCCTTCTTCCCTTAAC 1883
 QY 1796 TTTCCTCTCTCTCTCTGACCCCTTCTTCTTCTTCTTCTGCTGCTGGAAGTTATCTTCT 1855
 DB 1884 TTTCCTCTCTCTCTCTGACCCCTTCTTCTTCTTCTTCTGCTGCTGGAAGTTATCTTCT 1943
 QY 1856 GTGATTTGGTTCCTCTCTCTCCATATTTCTCCGCAATCATTAAGCTTCTTCTGA 1915
 DB 1944 GTGATTTGGTTCCTCTCTCTCCATATTTCTCCGCAATCATTAAGCTTCTTCTGA 2003
 QY 1916 GCTCAGCTGCTTCTTAATGGGGAATCCCTGCAAAACAAAGCTGCTTGGCTGATTTTATGG 1975
 DB 2004 GCTCAGCTGCTTCTTAATGGGGAATCCCTGCAAAACAAAGCTGCTTGGCTGATTTTATGG 2063
 QY 1976 ACTTAGAATGATCCCTTAACAGATTCCTTTCACATCAAAAAGTGCATTTGTCTGGAA 2035
 DB 2064 ACTTAGAATGATCCCTTAACAGATTCCTTTCACATCAAAAAGTGCATTTGTCTGGAA 2123
 QY 2036 GGCAACAGATTTCTTCAATTAATTGGAAAGGCTGCGCTATTAGTTGTCATTAAT 2095
 DB 2124 GGCAACAGATTTCTTCAATTAATTGGAAAGGCTGCGCTATTAGTTGTCATTAAT 2183
 QY 2096 GGTTTTGTACATCATTAATTAATCAAGATGTAATAGGAAAAA 2144
 DB 2184 GGTTTTGTACATCATTAATTAATCAAGATGTAATAGGAAAAA 2232

RESULT 4
 AX746724 2092 bp mRNA linear PAT 20-JUN-2003
 DEFINITION Sequence 249 from Patent EP1308459.
 ACCESSION AX746724
 VERSION AX746724.1 GI:32130991
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Isozaki, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Iseino, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamachi, K. I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahashi, K. and Masuhio, Y.
 TITLE Full-length cDNA sequences
 JOURNAL Patent: EP 1308459-A 249 07-MAY-2003; Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
 FEATURES

source
 1..2092
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 88.8%; Score 1917; DB 6; Length 2092;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1940; Conservative 0; Mismatches 5; Indels 10; Gaps 1;

QY 186 GGGCTTCTGTCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCAA 245
 DB 138 GGGCTTCTGTCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCAA 197
 QY 246 GTCACTGTCTGGAAAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCCCTTGTGCT 305
 DB 198 GTCACTGTCTGGAAAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCCCTTGTGCT 257
 QY 306 CCCCCGCTGGGATCCCTGATTTTGAACCCGAGCCCTGCATCCGCGCTGGATTAATTG 365
 DB 258 CCCCCGCTGGGATCCCTGATTTTGAACCCGAGCCCTGCATCCGCGCTGGATTAATTG 317
 QY 366 CGAAGAGCCACTCTTACCTTAATTT-----GTGCTTCAAGACTCAGAGTGGCT 415
 DB 318 CGAAGAGCCACTCTTACCTTAATTTGTAAGACAGAGTGGCTTCAAGACTCAGAGTGGCT 377
 QY 416 GCTTTTGAATCAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGCTG 475
 DB 378 GCTTTTGAATCAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGCTG 437
 QY 476 CCTTACCTGAACATCTAATGGGCTGGCCAGCCGATTAAGGCTCCAGCTCCCGCTT 535
 DB 438 CCTTACCTGAACATCTAATGGGCTGGCCAGCCGATTAAGGCTCCAGCTCCCGCTT 497
 QY 536 GGTGTGTTCCTCAGAGAGGTGCTTCAAGACTGCTCAAGCTTCATCTTTGATGGTCCG 595
 DB 498 GGTGTGTTCCTCAGAGAGGTGCTTCAAGACTGCTCAAGCTTCATCTTTGATGGTCCG 557
 QY 596 CCTGCTGCTTATGAGAGCTGCTGCTGTGTGTGCTGCTCAATCCGCTTGAATATTGG 655
 DB 558 CCTGCTGCTTATGAGAGCTGCTGCTGTGTGTGCTGCTCAATCCGCTTGAATATTGG 617
 QY 656 TTTCTTCCACCAATGGGATCAACATGCTCCGGGGAATCGGGCTTCAAGAGCAAGTGG 715
 DB 618 TTTCTTCCACCAATGGGATCAACATGCTCCGGGGAATCGGGCTTCAAGAGCAAGTGG 677
 QY 716 TGCTCTGTCTGGGCTTCAAGAAACATCGAGTTCTTGGTGGGACCCCAAGCTGTGAC 775
 DB 678 TGCTCTGTCTGGGCTTCAAGAAACATCGAGTTCTTGGTGGGACCCCAAGCTGTGAC 737
 QY 776 CATCTTTGGAGATCCGCGGAGCCATTAAGTGTCTTACTTTATCTGTCTCCATGGC 835
 DB 738 CATCTTTGGAGATCCGCGGAGCCATTAAGTGTCTTACTTTATCTGTCTCCATGGC 797
 QY 836 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGGGCTCATCATCTTAACTCTGA 895
 DB 798 CAAAGCTTATTCACAAAGCCATCATGAGAGTGGGGGCTCATCATCTTAACTCTGA 857
 QY 896 GGCCCATGATTAAGAGAGTGAAGACTCTGCAAGTGTGACATTTCTGTGTAAACA 955
 DB 858 GGCCCATGATTAAGAGAGTGAAGACTCTGCAAGTGTGACATTTCTGTGTAAACA 917
 QY 956 TGCGTCAAGCTGTGAGGCTCTGCTGAGTGTGAGGACAAACCTTCAAGAGTGTCT 1015
 DB 918 TGCGTCAAGCTGTGAGGCTCTGCTGAGTGTGAGGACAAACCTTCAAGAGTGTCT 977
 QY 1016 GACCTTCAAGCCAGAAACAAAGTCTTCTTCACTGAGTGTGATGAGTCTTCTTCTAA 1075
 DB 978 GACCTTCAAGCCAGAAACAAAGTCTTCTTCACTGAGTGTGATGAGTCTTCTTCTAA 1037
 QY 1076 TGAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT 1135
 DB 1038 TGAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGT 1097
 QY 1136 CATTAACAAGAGTGTGCTTCTGTGCTTAAGAGAGGCTCTGAGATCTCATGTG 1195
 DB 1098 CATTAACAAGAGTGTGCTTCTGTGCTTAAGAGAGGCTCTGAGATCTCATGTG 1157
 QY 1196 CTCACAAGATCCCTTGTGCTTCAATCTGATTAACAAACATCTGCAATCCCGCTCAGTA 1255
 DB 1158 CTCACAAGATCCCTTGTGCTTCAATCTGATTAACAAACATCTGCAATCCCGCTCAGTA 1217
 QY 1256 TTGCACTTGTGGCTAATGAATACCTTCATGACAAAGCACTCCCTGACTGAATCCGAGA 1315

```

Db      1218  TTTGCACTTGGCTAATGAAATCTTCAATGACAGACAGCTCCCTGACTGAAATCCAGAA 1277
Oy      1316  CAGCTTCTGCACTTGGCTGAGATGTTCTTGTGGTCCCTGCACTGATCAGAGCTCG 1375
Db      1278  CAGCTTCTGCACTTGGCTGAGATGTTCTTGTGGTCCCTGCACTGATCAGAGCTCG 1337
Oy      1376  ATATACAGAGATGCTGGTGCACCGTCTACTCTTAATGAGTTTGGGACCGGCTCTGAG 1435
Db      1338  ATATACAGAGATGCTGGTGCACCGTCTACTCTTAATGAGTTTGGGACCGGCTCTGAG 1397
Oy      1436  CTTTGAAGACAGGAGCGGCTTTTGTCAAGCCGACCACTGATGAGTCCGCTTGT 1495
Db      1398  CTTTGAAGACAGGAGCGGCTTTTGTCAAGCCGACCACTGATGAGTCCGCTTGT 1457
Oy      1496  GTTCGGTGGTCCCTTCCGAAAGGGGGGCAATTGTAATTTGGAAGAGCCGAGAGAGA 1555
Db      1458  GTTCGGTGGTCCCTTCCGAAAGGGGGGCAATTGTAATTTGGAAGAGCCGAGAGAGA 1517
Oy      1556  GAAGTTACTAGCGCGAAGATGATGAAATCTGGGCTACCTTGTCTGCAACCGGAATCC 1615
Db      1518  GAAGTTACTAGCGCGAAGATGATGAAATCTGGGCTACCTTGTCTGCAACCGGAATCC 1577
Oy      1616  TAAATGGGAAGAGCTGTCTGTGGCCAGCTTAAATCTGACTGAGAGTACTCTCAGCT 1675
Db      1578  TAAATGGGAAGAGCTGTCTGTGGCCAGCTTAAATCTGACTGAGAGTACTCTCAGCT 1637
Oy      1676  GAACTTGAACATGAGCTCGGACAGAGACTCAAGAAACCGGGGAGATTTTGGACAG 1735
Db      1638  GAACTTGAACATGAGCTCGGACAGAGACTCAAGAAACCGGGGAGATTTTGGACAG 1697
Oy      1736  CAGCAATCCCTGATCTGTCTGCTCCGACATGCTCCAAGTCTCTTCTTCTTAAAC 1795
Db      1698  CAGCAATCCCTGATCTGTCTGCTCCGACATGCTCCAAGTCTCTTCTTCTTAAAC 1757
Oy      1796  TTTCTCTCTCTCTCTCAGGCTTTTCTTCTTGTGCTCTTGAAGATTAATCTTCT 1855
Db      1758  TTTCTCTCTCTCTCTCAGGCTTTTCTTCTTGTGCTCTTGAAGATTAATCTTCT 1817
Oy      1856  GTGATTTTGGTTCCTTCTCTCCCAATATTTCTCCGCAATCAATAGCTTCTTCTGA 1915
Db      1818  GTGATTTTGGTTCCTTCTCTCCCAATATTTCTCCGCAATCAATAGCTTCTTCTGA 1877
Oy      1916  GCTCAGCTGCTTCTTAATGGGGATCTTGAAGAAACAGCTGCTTGGCTGATATTTATG 1975
Db      1878  GCTCAGCTGCTTCTTAATGGGGATCTTGAAGAAACAGCTGCTTGGCTGATATTTATG 1937
Oy      1976  ACTTAGAATGATCTCTTACAGAAATCTTTTCAACATCAAAAAGTGCATTTGTCTGAA 2035
Db      1938  ACTTAGAATGATCTCTTACAGAAATCTTTTCAACATCAAAAAGTGCATTTGTCTGAA 1997
Oy      2036  GGCAACAGATTTCTTCAATTAATTTGGAAGAGGCTGCTTATTAAGTTGTCATTAAT 2095
Db      1998  GGCAACAGATTTCTTCAATTAATTTGGAAGAGGCTGCTTATTAAGTTGTCATTAAT 2057
Oy      2096  GGTTTTGTACTCATATGAATTAATCAAGATG 2130
Db      2058  GGTTTTGTACTCATATGAATTAATCAAGATG 2092

```

```

RESULT 5
AK090997 2092 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ33678 file clone BRAHM2002191, weakly similar
DEFINITION AK090997 to FATTY ACYL-CoA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).
ACCESSION AK090997.1 GI:21749268
VERSION oligo capping; fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

```

AUTHORS

Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Nakita, H., Sekine, M., Obayashi, M., Niishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosokiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Fujiyama, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murokawa, K., Fujimori, K., Tanai, H., Kimura, M., Watanabe, M., Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoshida, M., Houta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiyama, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mutsaers, K., Yuuki, F., Hara, R., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Saich, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kunigai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tanihara, H., Tanigami, A., Fujizawa, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL PUBLISHED

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

COMMENT

FEATURES

source

CDS

```

1. 2092
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAHM2002191"
/tissue_type="brain"
/clone_lib="BRAHM2"
/note="Cloning vector: pME18SFL3"
394..1803
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC03565.1"
/db_xref="GI:21749268"
/translation="MLKHYHFKRVSDECLYATLYPAHADDTGSKLPLVWVPGAGAK
TGSATIPGSAALAAEDLVVVOYRIGLIGFETTWQHAFGNNAFQDVAAALSWQK

```


NIEPFGDPSVITFGESGAGIAVSLLISLSPAKGLFKHAKMESGVAIIPLYLEAHDE
 KSEBIOVAHAFKCNNA SDSBALRLCRFLKPSKSLTFSKRSFPRVVDGAFPEPPL
 DLISOKAFAPISIIIVNSHSCGFLIPKHEAPILLSGNSKSLAHLIONIIHIPROYL
 HUVANEYPHDHSLSLREIRDSLLDLGIDVFPVVPALITRHYRDNAGAPYTFEPRRPO
 CFEDTKPAFKADHADVRFVFGAFNLGDIVMEGATEEBELLSRKMKWATPART
 GPNPNDLSLWPAVNLBQYIOLDLMSLGGRLKRPVEFWTSTIPLILASDMHSP
 LSLITLISLQPPFFFCAP"

ORIGIN

Query Match 88.8%; Score 1917; DB 9; Length 2092;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1940; Conservative 0; Mismatches 5; Indels 10; Gaps 1;

186 GGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGTGATTCAGGGCAAGCAA 245
 138 GGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGTGATTCAGGGCAAGCAA 197
 246 GTCACTGTGTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTTGCTGCT 305
 198 GTCACTGTGTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTTGCTGCT 257
 306 CCCCCGCTGGATCCCTGCGATTTACGAACCGCGAGCTGCATCCCTGGGATTAATTG 365
 258 CCCCCGCTGGATCCCTGCGATTTACGAACCGCGAGCTGCATCCCTGGGATTAATTG 317
 366 CGAGAGCAGCCTCCTCACTTAATTT-----GTGCTCCAGAGCTCGAGAGTGGCT 415
 318 CGAGAGCAGCCTCCTCACTTAATTTGTGAACAGAGGCTCCCGAGAGCTCGAGAGTGGCT 377
 416 GCTCTTAAGTCAACATGCTCAAGTGCATTTACCGGAATTCGAGTGTGAGAGAGCTG 475
 378 GCTCTTAAGTCAACATGCTCAAGTGCATTTACCGGAATTCGAGTGTGAGAGAGCTG 437
 476 CCTCTAAGTCAACATGCTCAAGTGCCTCCAGCGATACAGGCTTCAAGTCTCCCTGCTT 535
 438 CCTCTAAGTCAACATGCTCAAGTGCCTCCAGCGATACAGGCTTCAAGTCTCCCTGCTT 497
 536 GGTGTGTTCCAGAGAGGCTTCAAGTGCCTCCAGCGATACAGGCTTCAAGTGCCTCCG 595
 498 GGTGTGTTCCAGAGAGGCTTCAAGTGCCTCCAGCGATACAGGCTTCAAGTGCCTCCG 557
 596 CCTGTGCTCATGAGAGAGCTGTGCTGTGTGTGCTGTGCTCAAGTACCGGCTAGGAATTTTGG 655
 558 CTTGGCTGCTATAGAGAGCTGTGCTGTGTGTGCTGTGCTCAAGTACCGGCTAGGAATTTTGG 617
 656 TTTCTTCAACATGAGGATCAAGTGTCTCCGGGAACTGGGCTTCAAGAGCAAGGTGGC 715
 618 TTTCTTCAACATGAGGATCAAGTGTCTCCGGGAACTGGGCTTCAAGAGCAAGGTGGC 677
 716 TGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 775
 678 TGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 737
 776 CATCTTTGGCGAGTCCGCGGAGCAATAAGTGTCTTCAAGTGTGCTGTGCTGTGCTGTGCT 835
 738 CATCTTTGGCGAGTCCGCGGAGCAATAAGTGTCTTCAAGTGTGCTGTGCTGTGCTGTGCT 797
 836 CAAAGGCTTAATTCACAAAGCATCATGAGAGTGGGGTGGCCATATCCCTTACCTTGA 895
 798 CAAAGGCTTAATTCACAAAGCATCATGAGAGTGGGGTGGCCATATCCCTTACCTTGA 857
 896 GGGCCATGATTAATGAGAGAGTGAAGAGCTTCAAGTGTGCTGTGCTGTGCTGTGCTGTGCT 955
 858 GGGCCATGATTAATGAGAGAGTGAAGAGCTTCAAGTGTGCTGTGCTGTGCTGTGCTGTGCT 917
 956 TGGCTCAAGTCTGAGGCTCTGTGAGGCTGTGCTGTGAGCAAAACCTTCAAGAGAGCTCT 1015
 918 TGGCTCAAGTCTGAGGCTCTGTGAGGCTGTGCTGTGAGCAAAACCTTCAAGAGAGCTCT 977
 1016 GACCTTCAGCAAGAAACAAAGTCTTTCATCTGAGTGTGAGTGTGCTTCTTCTTAA 1075
 978 GACCTTCAGCAAGAAACAAAGTCTTTCATCTGAGTGTGAGTGTGCTTCTTCTTAA 1037

1076 TGAGCTCTAAGTCTAATGCTCTGAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGT 1135
 1038 TGAAGCTCTAAGTCTAATGCTCTGAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGT 1097
 1136 CAATTAACAGAGTGTGCTCTGCTGCTGCTATGAGAGAGGCTCTGAGATCTTCAGTGG 1195
 1098 CAATGACAGAGTGTGCTCTGCTGCTGCTATGAGAGAGGCTCTGAGATCTTCAGTGG 1157
 1196 CTCCAAAGAGTCCCTGCGCTTCATCTGATCAAAACATCTCCGCAATCCCGGCTCAATA 1255
 1158 CTCCAAAGAGTCCCTGCGCTTCATCTGATCAAAACATCTCCGCAATCCCGGCTCAATA 1217
 1256 TTTGACCTTGTGCTAATGATACTTCATGACAGAGCACTCCCTGACTGAATCCGAGA 1315
 1218 TTTGACCTTGTGCTAATGATACTTCATGACAGAGCACTCCCTGACTGAATCCGAGA 1277
 1316 CAGTCTCTGAGCTTGTGAGAGATGTGTTCTTTGTGCTCCCTGCACTGATCAAGCTGG 1375
 1278 CAGTCTCTGAGCTTGTGAGAGATGTGTTCTTTGTGCTCCCTGCACTGATCAAGCTGG 1337
 1376 ATATCAAGAGATGCTGTGAGACCTGTCTACTATGAGATTTGCGACCGGCTCAAGTGG 1435
 1338 ATATCAAGAGATGCTGTGAGACCTGTCTACTATGAGATTTGCGACCGGCTCAAGTGG 1397
 1436 CTTTGAAGACAGAGCCGGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1495
 1398 CTTTGAAGACAGAGCCGGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1457
 1496 GTTGGTGTGCTTCTGAGAGGGGAGCAATTTATGTTTCAAGAGGCTCAGCGAGAGAGA 1555
 1458 GTTGGTGTGCTTCTGAGAGGGGAGCAATTTATGTTTCAAGAGGCTCAGCGAGAGAGA 1517
 1556 GAAGTTCAGAGCCGGAGAGATGAATACTGGGCTACTTGTGCTGCAACCGGAAATCC 1615
 1518 GAAGTTCAGAGCCGGAGAGATGAATACTGGGCTACTTGTGCTGCAACCGGAAATCC 1577
 1616 TAATGGAGACAGCTGTCTGTGAGCCAGTTATATGATGATGAGAGAGTACTTCACT 1675
 1578 TAATGGAGACAGCTGTCTGTGAGCCAGTTATATGATGATGAGAGAGTACTTCACT 1637
 1676 GGACTTGAACATGAGCTCTGAGACAGAGACTCAAAAGAACCGGCGGTGATTTTGTGAGCAG 1735
 1638 GGACTTGAACATGAGCTCTGAGACAGAGACTCAAAAGAACCGGCGGTGATTTTGTGAGCAG 1697
 1736 CACCATCCCTGATACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795
 1698 CACCATCCCTGATACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757
 1796 TTTCT 1855
 1758 TTTCT 1817
 1856 GTGATTTGTGCT 1915
 1818 GTGATTTGTGCT 1877
 1916 GCTGAGCTGCTTTCTAATGAGGAGTCTTGCAGAAACAGAGTGTGCTGCTGCTGCTGCTGCT 1975
 1878 GCTGAGCTGCTTTCTAATGAGGAGTCTTGCAGAAACAGAGTGTGCTGCTGCTGCTGCTGCT 1937
 1976 ACTTAGAGATGATCTTACAGAAATTTCTTCAACATCAAAAGTGCATTTGTCTTGA 2035
 1938 ACTTAGAGATGATCTTACAGAAATTTCTTCAACATCAAAAGTGCATTTGTCTTGA 1997
 2036 GGCACAGAGATTTCTTCAATTAATTTGAGAGAGGCTGCTATTAATGATTAAT 2095
 1998 GGCACAGAGATTTCTTCAATTAATTTGAGAGAGGCTGCTATTAATGATTAAT 2057
 2096 GGTTTTGAATCTCATGAGAAATTAATCAAGATGT 2130
 2058 GGTTTTGAATCTCATGAGAAATTAATCAAGATGT 2092

RESULT 6
BC069501
DEFINITION Homo sapiens hypothetical protein FLJ31547, mRNA (cDNA clone
MGC:97182 IMAGE:7262427), complete cds.
ACCESSION BC069501
VERSION BC069501.1 GI:46854682
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2047)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Saeed, A.I., White, O.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stephenson, M., Soares, M.B., Donald, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McGwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Schevchenko, Y.,
Boutard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skleka, U., Small, D.B.,
Schreier, A., Schein, J.E., Jones, S.J. and Marra, M.A.
human and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2047)
DIRECTOR MGC Project.
TITLE Direct Submission
AUTHORS Submitted (29-APR-2004) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Baylor Human Genome Sequencing Center
CDNA Library Preparation: Baylor Human Genome Sequencing Center
DNA Sequencing By: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hilyk, S.W., Louised, H.,
Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navevati,
A.N., Gibbs, R.A.
FEATURES
Source
FEATURES
Source
gene
CDS
121..1848

/gene="FLJ31547"
/codon_start=1
/product="FLJ31547 protein"
/protein_id="AAH69501.1"
/db_xref="GI:46854683"
/db_xref="LOCUSID:221223"
/translation="MSGWVHPGQIILWALVTLAFLPKPSABGPQNTLGIWIGKQ
VTVLGSFVFNVPFVAPPLSLRFTNPQPSFMDNLRATSYENLCLIONSEML
LDHMLKVPKPFVSEBDCILNTYAPADHTGSKLPVLWFPQGAFTGASIFDGS
ALAAVEDLVVVVQYRIGIFGFTTMDQHPAGNMAFKDOVAALSMVKNTFEGDPS
SVTIFGESAGAISSVSLISPMAGLPHKAIMESGVALITYLEHVDKEDQVNH
FCGNNAISPSEALRLKRTKDSKELTTSOKTSPTRVVDCAFPENEBLDSQAFRA
IPSTIGYNNHEGCFLLPKKAPRILSSNSIALHLQNTLHPQYTLVAVEYFD
KHSITRDSLDLIDGVFFVPLVITARYHRAGAVFYEEPRHRPQCEPDYKPAV
KADHADRFRVFGGAFKGIIVMEGATEBKLSRKMVMYATFATGPNQNDLSL
WPAVNLTEOYLQDLNLSLQRLKEPRVDWTSTIPLLSASDMLHSPLSLFLSL
QPFPPFCAP"

ORIGIN

Query Match 86.0%; Score 1855; DB 9; Length 2047;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1855; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
186 GGGCTTCTGCTGAAGGAGCCACAGAGAACACGAGCTGGATGATTCAGGGCAAGCA 245
193 GGGCTTCTGCTGAAGGAGCCACAGAGAACACGAGCTGGATGATTCAGGGCAAGCA 252
246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGACGTGTCTCTGGAGTCCCTTGTGCTGT 305
253 GTCACTGTGCTGGAGAGCCCTGTGCTGTGACGTGTCTCTGGAGTCCCTTGTGCTGT 312
306 CCCCCGTGGGATCCCTGCGATTACGAACCCGACGCTGCATGCGCCGGGATACCTTG 365
313 CCCCCGTGGGATCCCTGCGATTACGAACCCGACGCTGCATGCGCCGGGATACCTTG 372
366 CGAAGACCACTCTCTCACTTAATTTGTGCTTCAGAACTCAGAGTGGCTCTTAGAT 425
373 CGAAGACCACTCTCTCACTTAATTTGTGCTTCAGAACTCAGAGTGGCTCTTAGAT 432
426 CAACACATGCTCAAGGAGGATTAACCCGAAATTCGAGGTGTCAGAAAGTGCCTTACCTG 485
433 CAACACATGCTCAAGGAGGATTAACCCGAAATTCGAGGTGTCAGAAAGTGCCTTACCTG 492
486 AACATCTATGCGCTCTCCACGCGCATACAGGCTCCAACTCCCGCTTGGTGTGCTTC 545
493 AACATCTATGCGCTCTCCACGCGCATACAGGCTCCAACTCCCGCTTGGTGTGCTTC 552
546 CCGAGAGTGCCTTCAAGCTGAGCTCAAGCTCAATCTTGTATGGTCCGCTGCTGCTGC 605
553 CCGAGAGTGCCTTCAAGCTGAGCTCAAGCTCAATCTTGTATGGTCCGCTGCTGCTGC 612
606 TATGAGAGCGTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
613 TATGAGAGCGTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
666 ACATGGGATCAGCATGCTCCGAGGAACTGAGGCTTCAAGAACACAGTGGTCTGCTGCTC 725
673 ACATGGGATCAGCATGCTCCGAGGAACTGAGGCTTCAAGAACACAGTGGTCTGCTGCTC 732
726 TGGGTCAGAAAGATGAGGTTCTTGGTGGGAGCCCACTCTGTGACCATCTTTGGC 785
733 TGGGTCAGAAAGATGAGGTTCTTGGTGGGAGCCCACTCTGTGACCATCTTTGGC 792
786 GAGTCGCGGAGGACATAGTCTTCTAGTCTTATATGCTCTCCATGGCCAAAGGCTTA 845
793 GAGTCGCGGAGGACATAGTCTTCTAGTCTTATATGCTCTCCATGGCCAAAGGCTTA 852
846 TTCGCAAAAGCATATATGAGAGTGGGTCGATCATCTTCACTTGGAGGCCCATGAT 905
853 TTCGCAAAAGCATATATGAGAGTGGGTCGATCATCTTCACTTGGAGGCCCATGAT 912
906 TATGAGAGAGTGAAGACCTGAGGTGCTGACATTTCTGTGTAACATGCTGAC 965

Db 913 TATGAGAGAGTGAAGACCTGACAGTGTGTCACATTTCTGTGTGAACAATCGCTGACAG 972
 Qy 966 TCTGAGGCCCTGCTGAGTGGCCCTGAGAGAAAACCTCCAGAGAGCTGCTGACCCCTCAGC 1025
 Db 973 TCTGAGGCCCTGCTGAGTGGCCCTGAGAGAAAACCTCCAGAGAGCTGCTGACCCCTCAGC 1032
 Qy 1026 CAGAAAACAAAGTCTTTCACCTGAGAGTGTGATGTGCTTCTTCTTAATAGAGCTCTTA 1085
 Db 1033 CAGAAAACAAAGTCTTTCACCTGAGAGTGTGATGTGCTTCTTCTTAATAGAGCTCTTA 1092
 Qy 1086 GATCTATTGTCTCAGAAACATTTAAAGCAATCTTCCATCATCTGAGAGTCAATTAACAC 1145
 Db 1093 GATCTATTGTCTCAGAAACATTTAAAGCAATCTTCCATCATCTGAGAGTCAATTAACAC 1152
 Qy 1146 GAGTGTGCTTCTGCTGCTGCTTAAGAGAGGCTCCGAGATCTCTGAGGCTCCCAACAG 1205
 Db 1153 GAGTGTGCTTCTGCTGCTGCTTAAGAGAGGCTCCGAGATCTCTGAGGCTCCCAACAG 1212
 Qy 1206 TCCCTTGCCCTCATCTGATACAAAACATCTGACATCCCGCTCAGATTTTGACCTT 1265
 Db 1213 TCCCTTGCCCTCATCTGATACAAAACATCTGACATCCCGCTCAGATTTTGACCTT 1272
 Qy 1266 GTGGCTAATGAATTACTTCATGACAGACATCTCCCTGACTGAATCCGAGACAGTCTTCTG 1325
 Db 1273 GTGGCTAATGAATTACTTCATGACAGACATCTCCCTGACTGAATCCGAGACAGTCTTCTG 1332
 Qy 1326 GACTTGCTTGAAGATGTGTTCTTGTGTGTCCTGACCTGATCAAGCTTGATTCACAA 1385
 Db 1333 GACTTGCTTGAAGATGTGTTCTTGTGTGTCCTGACCTGATCAAGCTTGATTCACAA 1392
 Qy 1386 GATGCTGTCGACCTGTCTACTTCTTAAGATTTCGACACCGGCTCAGTCTTGAAGAC 1445
 Db 1393 GATGCTGTCGACCTGTCTACTTCTTAAGATTTCGACACCGGCTCAGTCTTGAAGAC 1452
 Qy 1446 AGGAGGCCGCTTGTGTCAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCTGCTG 1505
 Db 1453 AGGAGGCCGCTTGTGTCAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCTGCTG 1512
 Qy 1506 GCTTCTCTGAAGGGGACATTTGTATGTTTGAAGAGCCAGAGAGAGAGAAAGTTACTG 1565
 Db 1513 GCTTCTCTGAAGGGGACATTTGTATGTTTGAAGAGCCAGAGAGAGAGAAAGTTACTG 1572
 Qy 1566 AGCCGGAAGATGATGAATTAAGTGGCTACTTGTGCTGAACCGGGAATCTTAATGGAAAC 1625
 Db 1573 AGCCGGAAGATGATGAATTAAGTGGCTACTTGTGCTGAACCGGGAATCTTAATGGAAAC 1632
 Qy 1626 GACTGTCTCTGTGGCCAGCTTAATTAATCTGACTGAGCAGTACTCCAGCTGGAATTGAAC 1685
 Db 1633 GACTGTCTCTGTGGCCAGCTTAATTAATCTGACTGAGCAGTACTCCAGCTGGAATTGAAC 1692
 Qy 1686 ATGAGCTCTGAGACAGATCTCAAGAAACCGCGGGTGGATTTTGAACGACCATCCCC 1745
 Db 1693 ATGAGCTCTGAGACAGATCTCAAGAAACCGCGGGTGGATTTTGAACGACCATCCCC 1752
 Qy 1746 CTGATCTCTGTGCTCCGACAGATGCTCAAGTCTCTTCTCTTAATCTTCCCTCT 1805
 Db 1753 CTGATCTCTGTGCTCCGACAGTCTCAAGTCTCTTCTCTTAATCTTCCCTCT 1812
 Qy 1806 CTCTCTCAGACCTTCTTCTTCTTGTGTGCTCTTGAAGATTATCTTCTGTGATTTTG 1865
 Db 1813 CTCTCTCAGACCTTCTTCTTCTTGTGTGCTCTTGAAGATTATCTTCTGTGATTTTG 1872
 Qy 1866 TTTCCCTTCTCTCCCAATTAATTTCTCCGCAATCATTAAGTCTTCTGAGCTCAGCTGC 1925
 Db 1873 TTTCCCTTCTCTCCCAATTAATTTCTCCGCAATCATTAAGTCTTCTGAGCTCAGCTGC 1932
 Qy 1926 TTTCTATGAGGAAATCTTGAAGAAACAAGCTTGTGCTGATTTTGAAGCTTAGAAT 1985
 Db 1933 TTTCTATGAGGAAATCTTGAAGAAACAAGCTTGTGCTGATTTTGAAGCTTAGAAT 1992
 Qy 1986 GATCTTATCAGAAATCTTTCACATCAAAAAGTGAATTTGTCTTGAAGGCA 2040
 Db 1993 GATCTTATCAGAAATCTTTCACATCAAAAAGTGAATTTGTCTTGAAGGCA 2047

RESULT 7
 BC039073
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Loulseg, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAX Plate: 51 Row: k Column: 18
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.
 FEATURES
 source
 Location/Qualifiers
 1..2298
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5174937"
 /tissue_type="Brain, Lung, Testis, adult, pooled whole"
 /clone_id="NIH_MGC_115"
 /lab_host="DH10B"
 /note="vector: PCMV-SPORT6"
 ORIGIN
 Query Match 84.6%; Score 1824.6; DB 9; Length 2298;
 Best Local Similarity 93.5%; Pred No. 0;
 Matches 1969; Conservative 0; Mismatches 4; Indels 133; Gaps 1;
 Qy 186 GGGCTTGTGCTGAAGGGCCAGAGAAACACAGAGCTGGATGATTCAGGCGAAGCA 245
 Db 103 GGGCTTGTGCTGAAGGGCCAGAGAAACACAGAGCTGGATGATTCAGGCGAAGCA 162
 Qy 246 GTTACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTTGTGCT 305
 Db 163 GTTACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTTCTCTGGAGTCCCTTTGTGCT 222
 Qy 306 CCCCGCTGGATCCCTGCGATTTTGAACCCGAGCCGTGATCGCCCTGGATTAATTG 365
 Db 223 CCCCGCTGGATCCCTGCGATTTTGAACCCGAGCCGTGATCGCCCTGGATTAATTG 282
 Qy 366 CGAGAGCCACTCTTACCTTAATTTGTGCTTCAAGAACTCAGAGTGGCTGCTTAAGT 425
 Db 283 CGAGAGCCACTCTTACCTTAATTTGTGCTTCAAGAACTCAGAGTGGCTGCTTAAGT 342

[illegible]

QY	1374	CGATATCAAGAGAAATGCTGTGACACTGTACTCTTAATGAGTTTGGCACCGGCCTCAG	1433
Db	1423	CGATATCAAGAGAAATGCTGTGACACTGTACTCTTAATGAGTTTGGCACCGGCCTCAG	1482
QY	1434	TGCTTTGAAAGACAGAGCCGGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTT	1493
Db	1483	TGCTTTGAAAGACAGAGCCGGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGCTTT	1542
QY	1494	GTGTTCCGGTGGTCCCTTCTCTAAGGGGGACATTTGTAATGTTTGAAGAGCCAGAGAG	1553
Db	1543	GTGTTCCGGTGGTCCCTTCTCTAAGGGGGACATTTGTAATGTTTGAAGAGCCAGAGAG	1602
QY	1554	GAGAAATTAAGTACAGCCGAGAGATGATGAATAACTGGGCTACCTTGGCTCGAACCGGGAA	1613
Db	1603	GAGAAATTAAGTACAGCCGAGAGATGATGAATAACTGGGCTACCTTGGCTCGAACCGGGAA	1662
QY	1614	CCTAAATGGGAAGCAGACTGTCTCTGTGGCAGCTTATAATGTAAGTGAAGTACCTCAG	1673
Db	1663	CCTAAATGGGAAGCAGACTGTCTCTGTGGCAGCTTATAATGTAAGTGAAGTACCTCAG	1722
QY	1674	CTGGACTTGAACATGAGGCTCGGACAGAGACTCAAAAGACCGGGGTGGATTTTGGACC	1733
Db	1723	CTGGACTTGAACATGAGGCTCGGACAGAGACTCAAAAGACCGGGGTGGATTTTGGACC	1782
QY	1734	AGCAACATCCCGCTGATCTGTCTGCGCTCGAGACATGCTCAAGTCCCTCTTCTTCTTA	1793
Db	1783	AGCAACATCCCGCTGATCTGTCTGCGCTCGAGACATGCTCAAGTCCCTCTTCTTCTTA	1842
QY	1794	ACTTTCCTCTCTCTCTCCAGCCTTTCTTTTCTTTTGTGCTCCTTGAGAAATTAATCTTT	1853
Db	1843	ACTTTCCTCTCTCTCTCCAGCCTTTCTTTTCTTTTGTGCTCCTTGAGAAATTAATCTTT	1902
QY	1854	CTGTGATTTTGGTTTCCCTTCTCCTCCCAATATTTCTCCCGCAATCAATTAAGCTTCTTCT	1913
Db	1903	CTGTGATTTTGGTTTCCCTTCTCCTCCCAATATTTCTCCCGCAATCAATTAAGCTTCTTCT	1962
QY	1914	GAGCTCAGCTGCTTCTATGGGAGTCCTTGCAAAAGCAGCTGCTTGGCGGATAATTTAT	1973
Db	1963	GAGCTCAGCTGCTTCTATGGGAGTCCTTGCAAAAGCAGCTGCTTGGCGGATAATTTAT	2022
QY	1974	GGACTTGGAAATGATCCTTAACAGAAATCTTTTCAACATCAAAAAGTCAATTTGTCTTGG	2033
Db	2023	GGACTTGGAAATGATCCTTAACAGAAATCTTTTCAACATCAAAAAGTCAATTTGTCTTGG	2082
QY	2034	AAGGCAACAAGATTTTCTTCAATTAATTTGGAGAAGGCTGGCCTATTAAGTGTCAATAA	2093
Db	2083	AAGGCAACAAGATTTTCTTCAATTAATTTGGAGAAGGCTGGCCTATTAAGTGTCAATAA	2142
QY	2094	ATGGTTTGTACTCATATGAATAAATAACAGATTAATAATAGGAAAAAATTTTTAAAAA	2153
Db	2143	ATGGTTTGTACTCATATGAATAAATAACAGATTAATAATAGGAAAAAATTTTTAAAAA	2202
QY	2154	AAAAA 2158	
Db	2203	AAAAA 2207	

RESULT 8				
LOCUS	AR438765			
DEFINITION	AR438765	1746 bp	DNA	linear
ACCESSION	Sequence 3	from patent US 6664091.		PAT 20-FEB-2004
VERSION	AR438765			
KEYWORDS	AR438765.1	GI:42663740		
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1746)			
TITLE	Curtis,R.A.J. and Silos-Santiago,I.			
JOURNAL	53010, a human carboxylesterase family member and uses thereof			
FEATURES	Patent: US 6664091-A 316-DEC-2003;			
	Location/Qualifiers			
source	1..1746			

ORIGIN

/organism="unknown"
/mol_type="genomic DNA"

Query Match 80.9%; Score 1746; DB 6; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

96 ATGCCACAGGAGCTTACTCATCTGCTTCAACAATGCTGCTTTTCTGATTTTCCAGCCC 155
1 ATGCCACAGGAGCTTACTCATCTGCTTCAACAATGCTGCTTTTCTGATTTTCCAGCCC 60
156 CTGTGTGGACACAGACAGTGGGAAAACTGGGCTTCTGTGAAGGGCCACAGAGAAC 215
61 CTGTGTGGACACAGACAGTGGGAAAACTGGGCTTCTGTGAAGGGCCACAGAGAAC 120
216 ACCAGGCTGGGATGATTCAGGGCAAGCACTGTGCTGGAAAGCCCTGTGCTGTG 275
121 ACCAGGCTGGGATGATTCAGGGCAAGCACTGTGCTGGAAAGCCCTGTGCTGTG 180
276 AACGTTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGGAATTTACGAAC 335
181 AACGTTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGGAATTTACGAAC 240
336 CCGACGCTGCATCCGCTGGGATTAATTCAGAGAGCCACTCTAATTTATTTGTC 395
241 CCGACGCTGCATCCGCTGGGATTAATTCAGAGAGCCACTCTAATTTATTTGTC 300
396 CTCAGAACTCAGAGTGGCTCTTAAATCAACAATGCTCAAGTGCATTAACCGAAA 455
301 CTCAGAACTCAGAGTGGCTCTTAAATCAACAATGCTCAAGTGCATTAACCGAAA 360
456 TTCCGAGTGCAGAGACTGCTCTAATCTGAACATTAAGGCGCTGCCAGCCGATACA 515
361 TTCCGAGTGCAGAGACTGCTCTAATCTGAACATTAAGGCGCTGCCAGCCGATACA 420
516 GAGCTCAAGCTCCCGCTGTGTGTGTTCCAGAGAGTCCCTCAAGAATGGCTCAGCC 575
421 GAGCTCAAGCTCCCGCTGTGTGTGTTCCAGAGAGTCCCTCAAGAATGGCTCAGCC 480
576 TCCATCTTTGATGGTCCGCTGCTGCTAATGAAGAGTGTGGTGTGCTCAG 635
481 TCCATCTTTGATGGTCCGCTGCTGCTAATGAAGAGTGTGGTGTGCTCAG 540
636 TACCGGCTAGAGAAATTTGGTCTTTCACACATGGGATCAGCATGCTCCGGGAACTGG 695
541 TACCGGCTAGAGAAATTTGGTCTTTCACACATGGGATCAGCATGCTCCGGGAACTGG 600
696 GCGTCAAGAGACAGATGGCTGCTGCTGCTGAGGATCAAGAAACATGAGTTCTTCGGT 755
601 GCGTCAAGAGACAGATGGCTGCTGCTGCTGAGGATCAAGAAACATGAGTTCTTCGGT 660
756 GGGGACCCAGCTCTGTGACATCTTTTGGCAGTCCGCGGAGCCATTAAGTCTTCAAT 815
661 GGGGACCCAGCTCTGTGACATCTTTTGGCAGTCCGCGGAGCCATTAAGTCTTCAAT 720
816 CTTAATCTGTCTCCATGAGCCAAAGCTTATTCACAAAGCATATGAGAGTGGGCTG 875
721 CTTAATCTGTCTCCATGAGCCAAAGCTTATTCACAAAGCATATGAGAGTGGGCTG 780
876 GGCATCATCCCTTACCTGAGGCCCAGATTAATGAAGAGTGAAGCCTGAGAGTGT 935
781 GGCATCATCCCTTACCTGAGGCCCAGATTAATGAAGAGTGAAGCCTGAGAGTGT 840
936 GCACATTTCTGTGTAACATGCGTCAAGCTGAGAGCCCTGCTGAGGTGCTGAGACA 995
841 GCACATTTCTGTGTAACATGCGTCAAGCTGAGAGCCCTGCTGAGGTGCTGAGACA 900
996 AAACCTTCAAGAGAGTGTGACCTCAGCCGAGAAAAAACAATCTTCACTGAGTGT 1055
901 AAACCTTCAAGAGAGTGTGACCTCAGCCGAGAAAAAACAATCTTCACTGAGTGT 960
1056 GATGTGCTTTCTTCTTAATGAGCCTTAATGATTAATGTCAGAAAGCATTTAAGCA 1115

|||||
961 GATGTGCTTTCTTCTTAATGAGCCTTAATGATTAATGTCAGAAAGCATTTAAGCA 1020
1116 ATTCCTTCAATCAGAGTGAATTAACAAGAGTGTGCTCCGCTGCTAATGAAGAG 1175
1021 ATTCCTTCAATCAGAGTGAATTAACAAGAGTGTGCTCCGCTGCTAATGAAGAG 1080
1176 GCTCCTGAGATCTCAGTGTGCTCAACAAGTCCCTGCTCATCTGATACAAACATC 1235
1081 GCTCCTGAGATCTCAGTGTGCTCAACAAGTCCCTGCTCATCTGATACAAACATC 1140
1236 CTGACATCCGCTCAGTATTTGACCTTGTGCTAATGAATCTTCATGACAGAC 1295
1141 CTGACATCCGCTCAGTATTTGACCTTGTGCTAATGAATCTTCATGACAGAC 1200
1236 TCCCTGATGAATTCGAGACAGTCTTGTGACCTTGTGAGAGTGTCTTGTGTC 1355
1201 TCCCTGATGAATTCGAGACAGTCTTGTGACCTTGTGAGAGTGTCTTGTGTC 1260
1356 CCGACATGATCAGAGTGAATTAACAAGAGTGTGCTGCTGCTAATCTTATGAG 1415
1261 CCGACATGATCAGAGTGAATTAACAAGAGTGTGCTGCTGCTAATCTTATGAG 1320
1416 TTTCGACACCGGCTCAGTGTGCTTGAAGACAGAAAGCCGCTTTGTCAAGCCGAC 1475
1321 TTTCGACACCGGCTCAGTGTGCTTGAAGACAGAAAGCCGCTTTGTCAAGCCGAC 1380
1476 GCTGATGAATTCGCTTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1535
1381 GCTGATGAATTCGCTTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
1536 GAAGAGGACAG 1595
1441 GAAGAGGACAG 1500
1596 TTTCGCTCAACCGGAGATCTTAATGGAACAGACCTGTCTGTGCGACAGCTTAATCTG 1655
1501 TTTCGCTCAACCGGAGATCTTAATGGAACAGACCTGTCTGTGCGACAGCTTAATCTG 1560
1656 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCTTCGACAGAGACTCAAAAGACG 1715
1561 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCTTCGACAGAGACTCAAAAGACG 1620
1716 GGGGTGAATTTTGAACAGACAGACATCCCGCTGATCCTGTGCTCCGACATGCTCAC 1775
1621 GGGGTGAATTTTGAACAGACAGACATCCCGCTGATCCTGTGCTCCGACATGCTCAC 1680
1776 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
1681 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1836 CCTTGA 1841
1741 CCTTGA 1746

RESULT 9
AX600126 1746 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 3 from Patent WO0250256.
ACCESSION AX600126
VERSION AX600126.1 GI:28400192
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Curtis, R.A. and Silos-Santiago, I.
53010, a novel human carboxylesterase family member and uses
thereof
Patent: WO 0250256-A 3 27-JUN-2002;
JOURNAL Millennium Pharmaceuticals, Inc. (US)

FEATURES
source 1. 1746
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 80.9%; Score 1746; DB 6; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 ATGCCACAGGAGCTTACTTCACTGCTTCACAATGAGTCTTTTCTGATTTCCAGCCC 155
Db 1 ATGCCACAGGAGCTTACTTCACTGCTTCACAATGAGTCTTTTCTGATTTCCAGCCC 60
QY 156 CTGTGGGGAACACAGACAGTGGGGAACCTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
Db 61 CTGTGGGGAACACAGACAGTGGGGAACCTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
QY 216 ACCAGGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAGCCCTGTGCTGTG 275
Db 121 ACCAGGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAGCCCTGTGCTGTG 180
QY 276 AAGTGTCTCTGGAAGTCCCTTTGCTGCTCCCGCTGGATCCCTGCAATTACGAAC 335
Db 181 AAGTGTCTCTGGAAGTCCCTTTGCTGCTCCCGCTGGATCCCTGCAATTACGAAC 240
QY 336 CCGCAGCTGCAATCGGCTGGGGAATACCTGCGAAGAACCACTCTCACTTAACTTTGTC 395
Db 241 CCGCAGCTGCAATCGGCTGGGGAATACCTGCGAAGAACCACTCTCACTTAACTTTGTC 300
QY 396 CTCCAGAACTCAGAGTGGCTGCTTATGATCAACATGCTCAAGGTGATTACCGGAA 455
Db 301 CTCCAGAACTCAGAGTGGCTGCTTATGATCAACATGCTCAAGGTGATTACCGGAA 360
QY 456 TTCCGAGTCTCAGAAAGTGGCTTAACTTACCTGAACATTTATGCGCTGCCAGCCATCA 515
Db 361 TTCCGAGTCTCAGAAAGTGGCTTAACTTACCTGAACATTTATGCGCTGCCAGCCATCA 420
QY 516 GGGCTCAAGCTCCCGCTTGTGATGATTCAGGAGGTGCTTCAAGATGAGCTCAAGC 575
Db 421 GGGCTCAAGCTCCCGCTTGTGATGATTCAGGAGGTGCTTCAAGATGAGCTCAAGC 480
QY 576 TCCATCTTGAATGGGTCCGCTGCTGCTATGAGAACGTGCTGCTGCTGCTGCTGCTG 635
Db 481 TCCATCTTGAATGGGTCCGCTGCTGCTATGAGAACGTGCTGCTGCTGCTGCTGCTG 540
QY 636 TACCGGCTAGGAATATTTGTTCTTCAACAATGGGATCAGCAATGCTCCGGGAACTGG 695
Db 541 TACCGGCTAGGAATATTTGTTCTTCAACAATGGGATCAGCAATGCTCCGGGAACTGG 600
QY 696 GCGTTCAAGAACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
Db 601 GCGTTCAAGAACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 756 GGGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 815
Db 661 GGGGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 816 CTATATCTGCTCCATGCGCAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTG 875
Db 721 CTATATCTGCTCCATGCGCAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTG 780
QY 876 GCGATCATCTCTTACCTGAGAGCCCATGATTTATGAGAAAGTGAAGACCTGAGAGTGT 935
Db 781 GCGATCATCTCTTACCTGAGAGCCCATGATTTATGAGAAAGTGAAGACCTGAGAGTGT 840
QY 936 GCACATTTCTGTGTAACAATGCGTCAAGCTGAGAGCCCTGCTGAGAGTCCGTAAGACA 995
Db 841 GCACATTTCTGTGTAACAATGCGTCAAGCTGAGAGCCCTGCTGAGAGTCCGTAAGACA 900
QY 996 AAACCTTCAAGAGTGTGACCTCAGCCAGAAAACAAAGTCTTCACTCGAGTGTGTT 1055

Db 901 AAACCTTCAAGAGTGTGACCTCAGCCAGAAAACAAAGTCTTCACTCGAGTGTG 960
QY 1056 GANGGGCTTTCTTCTCTTAATGAGGCTTAAGTCTATGCTCAAGAAACATTTAAAGCA 1115
Db 961 GANGGGCTTTCTTCTCTTAATGAGGCTTAAGTCTATGCTCAAGAAACATTTAAAGCA 1020
QY 1116 ATTCTTCATCATCGAGTCAATTAACACAGAGTGTGCTTCCGCTGCTCATGAGAGAG 1175
Db 1021 ATTCTTCATCATCGAGTCAATTAACACAGAGTGTGCTTCCGCTGCTCATGAGAGAG 1080
QY 1176 GCTCTGAGATTCCTAGTGGCTCCAAACAGTCCCTTGCCTCATCTGATACAAAACATC 1235
Db 1081 GCTCTGAGATTCCTAGTGGCTCCAAACAGTCCCTTGCCTCATCTGATACAAAACATC 1140
QY 1236 CTGCACATCCGCGCTCAGTATTTGACCTTGTGCTAATGAATCTTCCATGACACAC 1295
Db 1141 CTGCACATCCGCGCTCAGTATTTGACCTTGTGCTAATGAATCTTCCATGACACAC 1200
QY 1296 TCCCTGACTGAATCCGAGACAGTCTTCTGACTTGTGAGATGTGTTCTTTGTGCTG 1355
Db 1201 TCCCTGACTGAATCCGAGACAGTCTTCTGACTTGTGAGATGTGTTCTTTGTGCTG 1260
QY 1356 CCTGCACTGATCAGAGCTCCATATACAGAGATGCTGTGCACTGTCTATGATGAG 1415
Db 1261 CCTGCACTGATCAGAGCTCCATATACAGAGATGCTGTGCACTGTCTATGATGAG 1320
QY 1416 TTTGCGACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTTGTCAAAAGCCGAC 1475
Db 1321 TTTGCGACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTTGTCAAAAGCCGAC 1380
QY 1476 GCTGATGAATGCTGCTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1535
Db 1381 GCTGATGAATGCTGCTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1536 GAAGGACCCACGAGAGGAGAGATTACTAGGCGGAGATGATGAATATAGGGGTACC 1595
Db 1441 GAAGGACCCACGAGAGGAGAGATTACTAGGCGGAGATGATGAATATAGGGGTACC 1500
QY 1596 TTTGCTCGAAACCGGGAATCTTAATGGGAACAGCTGTCTGTGTGCGACCTTAATCTG 1655
Db 1501 TTTGCTCGAAACCGGGAATCTTAATGGGAACAGCTGTCTGTGTGCGACCTTAATCTG 1560
QY 1656 ACTGAGAGTAACTCCAGGCTGGACCTTGAACATGAGGCTGGACAGAGATCTCAAGAAC 1715
Db 1561 ACTGAGAGTAACTCCAGGCTGGACCTTGAACATGAGGCTGGACAGAGATCTCAAGAAC 1620
QY 1716 CCGGTGGAATTTTGAACAGACCATCCCTGATCTGTGCTGCTGCTGCTGCTGCTG 1775
Db 1621 CCGGTGGAATTTTGAACAGACCATCCCTGATCTGTGCTGCTGCTGCTGCTGCTG 1680
QY 1776 AGTCCCTTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
Db 1681 AGTCCCTTTCTTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
QY 1836 CCTTGA 1841
Db 1741 CCTTGA 1746

RESULT 10
AX714052
LOCUS AX714052 1962 bp. DNA linear PAT 15-APR-2003
DEFINITION Sequence 736 from Patent EP1293569.
ACCESSION AX714052
VERSION AX714052.1 GI:29888980
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS 1. Iwagaki, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J.I., Isono, Y., Hito, Y., Otsuka, K., Nagai, K., Irie, R.,

Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
Maehno, Y.
Full-length cDNAs
Patent: EP 1293569-A 736 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)

FEATURES
Source
Location/Qualifiers
1..1962
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 73.2%; Score 1579.4; DB 6; Length 1962;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;
186 GGGCCCTTGTGTAAGGGCCACAGAGAACACACAGGCTGGATTCAGGGCAAGCA 245
222 GGGCCCTTGTGTAAGGGCCACAGAGAACACACAGGCTGGATTCAGGGCAAGCA 281
246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTTGAGTCCCTTTGCTGT 305
282 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTTGAGTCCCTTTGCTGT 341
306 CCCCCGTGGGATCCCTGGAATTTAGAACCCGACGCTGATCGCCCTGGGATTAAGTTG 365
342 CCCCCGTGGGATCCCTGGAATTTAGAACCCGACGCTGATCGCCCTGGGATTAAGTTG 401
366 CGAGAGCCACCTCTTACCTTAATTTGTGCTTCAAGATCAGAGTGGCTCTTAAGT 425
402 CGAGAGCCACCTCTTACCTTAATTTGTGCTTCAAGATCAGAGTGGCTCTTAAGT 461
426 CAACCATGCTCAAGTGTGATTAACCCGAATTCGAGTGTCAAGAGCTGCTCTTAAGT 485
462 CAACCATGCTCAAGTGTGATTAACCCGAATTCGAGTGTCAAGAGCTGCTCTTAAGT 521
486 AACATCTATGCGCCCTGACCGACGATCAAGGCTCAAGCTCCCGTCTGGTGTGTTG 545
522 AACATCTATGCGCCCTGACCGACGATCAAGGCTCAAGCTCCCGTCTGGTGTGTTG 581
546 CCAGAGGTGCTTCAAGATCTGGCTCAAGCTTCAATCTTGTAGTGGTCCGCTTGCTGC 605
582 CCAGAGGTGCTTCAAGATCTGGCTCAAGCTTCAATCTTGTAGTGGTCCGCTTGCTGC 641
606 TATGAGAGCTGTGCTGTGTGTGCTCAAGTACCGGCTTAGAATTTGTTTCTTCAAC 665
642 TATGAGAGCTGTGCTGTGTGTGCTCAAGTACCGGCTTAGAATTTGTTTCTTCAAC 701
666 AACATCTATGCGCCCTGACCGACGATCAAGGCTCAAGCTCCCGTCTGGTGTGTTG 725
702 AACATCTATGCGCCCTGACCGACGATCAAGGCTCAAGCTCCCGTCTGGTGTGTTG 761
726 TGGGTCCAGAGAGATCGAGTTCTTGGTGGGAGCCCAAGCTCTGTGACATCTTTGGC 785
762 TGGGTCCAGAGAGATCGAGTTCTTGGTGGGAGCCCAAGCTCTGTGACATCTTTGGC 821
786 GAGTCCGCGGAGCCATTAAGTCTTCTAAGTCTTAATCTGTCTCCATGGCCAAAGCTTA 845
822 GAGTCCGCGGAGCCATTAAGTCTTCTAAGTCTTAATCTGTCTCCATGGCCAAAGCTTA 881
846 TTTCCCAAGAGCATATGAGAGAGTGGGGGATCAATCCCTTAAGTGGCCCAATAT 905
882 TTTCCCAAGAGCATATGAGAGAGTGGGGGATCAATCCCTTAAGTGGCCCAATAT 941
906 TATGAGAGAGTGAAGACCTGAGGTGTGTGACATTTCTGTGTAAACATCTGTCAGAC 965
942 TATGAGAGAGTGAAGACCTGAGGTGTGTGACATTTCTGTGTAAACATCTGTCAGAC 1001
966 TCTGAGGCTCTGTGAGGTGTCTGAGAGCAAAACCTTCAAGAGAGCTGTGACCTCAAC 1025
1002 TCTGAGGCTCTGTGAGGTGTCTGAGAGCAAAACCTTCAAGAGAGCTGTGACCTCAAC 1061

1026 CAGAAAACAAAGCTTTTCACTGAGAGTGTGATGAGTGTCTTTCTTCTAATGAGCTCTTA 1085
1062 CAGAAAACAAAGCTTTTCACTGAGAGTGTGATGAGTGTCTTTCTTCTAATGAGCTCTTA 1121
1086 GATCTATTGTCTCAGAAAGCATTTTAAAGCATTTCTTCAATCATGTGAGTCAATTAACAC 1145
1122 GATCTATTGTCTCAGAAAGCATTTTAAAGCATTTCTTCAATCATGTGAGTCAATTAACAC 1181
1146 GAGTGTGCTTCCGCTGCTATGAGAGAGGCTCTGAGATCTCAGTGTGCTCAACAG 1205
1182 GAGTGTGCTTCCGCTGCTATGAGAGAGGCTCTGAGATCTCAGTGTGCTCAACAG 1241
1206 TCCCTTGGCTTCCATCTGATCAAAACATCTGACATTCGCGCTCAAGTATTGACCTT 1265
1242 TCCCTTGGCTTCCATCTGATCAAAACATCTGACATTCGCGCTCAAGTATTGACCTT 1301
1266 GTGGCTAATGAATACTTCCATGACAGACATCTCCCTGACTGAAATCCGAGACAGTCTTCTG 1325
1302 GTGGCTAATGAATACTTCCATGACAGACATCTCCCTGACTGAAATCCGAGACAGTCTTCTG 1361
1326 GACTGTGCTGAGAGTGTCTTGTGTGCTCCCTGACATGATCAGAGTCTGATACAGA 1385
1362 GACTGTGCTGAGAGTGTCTTGTGTGCTCCCTGACATGATCAGAGTCTGATACAGA 1421
1386 GATGCTGTGACACTGTCTACTTATGATGATTTCCGACACCGGCTCAAGTCTTTGAAAGC 1445
1422 ----- 1421
1446 ACGAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGAAGTCCGTTGTTCGATGCT 1505
1422 ----- 1421
1506 GCTTCTCGAAGGGGACATTTGTTATGTTGGAAGAGCCAGGAGAGAGATTAAGT 1565
1422 -----GAGGAGCCAGAGAGAGAGATTAAGT 1451
1566 AGCCGGAAGATGATGAATAATCTGAGGCTACTTGTCTCAACCGGAAATCTTAATGGAAAC 1625
1452 AGCCGGAAGATGATGAATAATCTGAGGCTACTTGTCTCAACCGGAAATCTTAATGGAAAC 1511
1626 GACTGTCTGTGTGACAGCTTATATCTGACTGACAGTACCTTCAAGCTGACCTTGAAC 1685
1512 GACTGTCTGTGTGACAGCTTATATCTGACTGACAGTACCTTCAAGCTGACCTTGAAC 1571
1686 ATGAGGCTTCGAGCAGAGACTGAAGAAACCGGAGGAGATTTTGGACAGACCAATCCCC 1745
1572 ATGAGGCTTCGAGCAGAGACTGAAGAAACCGGAGGAGATTTTGGACAGACCAATCCCC 1631
1746 CTGATCTGTGTGCTCCGACATGCTCAGAGTCTCTTCTTCTTAATCTTCTCTCT 1805
1632 CTGATCTGTGTGCTCCGACATGCTCAGAGTCTCTTCTTCTTAACTTCTCTCT 1691
1806 CTCTCAGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1865
1692 CTCTCAGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1751
1866 TTTCCCTTCTCTCCCAAAATTTCTCCCGCAATCAATGCTTCTTCTTCTTCTTCTTCTT 1925
1752 TTTCCCTTCTCTCCCAAAATTTCTCCCGCAATCAATGCTTCTTCTTCTTCTTCTTCTT 1811
1926 TTTCTATGAGGATCTTTCAGAAACAAAGCTGCTTTCGATATTTTATGAGCTTGAAGT 1985
1812 TTTCTATGAGGATCTTTCAGAAACAAAGCTGCTTTCGATATTTTATGAGCTTGAAGT 1871
1986 GATCTTTCAGAAATCTTTCAGATCAATCAAAAGTGAATTTTGTCTTGAAGCAACAG 2045
1872 GATCTTTCAGAAATCTTTCAGATCAATCAAAAGTGAATTTTGTCTTGAAGCAACAG 1931
2046 TTTCTTCAATTAATTTGAGAGAGGCTGACC 2076
1932 TTTCTTCAATTAATTTGAGAGAGGCTGACC 1962

RESULT 11
AK056109
LOCUS
DEFINITION Homo sapiens cDNA FLJ31547 f1s, clone NT2R12001010, weakly similar to Fatty ACyl-CoA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).
ACCESSION AK056109
VERSION AK056109.1 GI:16551422
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Oka,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Iishi,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shitatori,A., Sudo,H., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ichibashi,T., Yamashita,H., Hiraoaka,S., Chiba,Y., Iehida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yoshida,M., Houta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanabe,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Muraashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotaka,S., Yoshikawa,Y., Matsumura,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mitoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hiroo,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Motinaga,M., Sasaki,M., Togaishi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Maehuo,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isegai,T. and Sugano,S.

TITLE
JOURNAL
PUBMED
AUTHORS
REFERENCE
14702039
Nat. Genet. 36 (1), 40-45 (2004)

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
COMMENT
Submitted (24-OCT-2001) Takao Isegai, Helix Research Institute, Genomics Laboratory, 1532-3 Yena, Kisarazu, Chiba 222-0812, Japan (E-mail:genomics@hi.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
Location/Qualifiers

source

1..1962
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2R12001010"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_11b="NT2R12"
/note="cloning vector: PMB18SF13-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction. -majorly NT2 neuron 150..1727
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71094.1"
/db_xref="GI:16551423"
/translation="MSGNWVHQIILTWAIWLAAPTKGSAAGPQNRNLGNIGQKQVTVGSPVPPVFPVGPAPRPPGSLAFETTPQASPDNDLREATSTPNLCLNSEWLLDQMLKXVHPREVSDCLYLNIVAPHADTGSKLPLVWPFQGAFFKGSASIFDGSALAAVEVLVAVVQVSLGIFFTTWDQNAQWMAFDQVVALSWQKNIFFGQDS SVTFEGSAGALISVSLILSPMAKGLFHKAINSGVALIPLYLAHYEKSEDOVVAH FCGNNAQDSBALRCLRTKSKELTSLQKTSFTFVVDGAFPPNPLDLSQAKRA IPSIIIVNHECGFLPEKPAPELISGNSKSLALHILIONLTIH PPOYHLVAEYEHDKSLTEIRDSLDLDLGDVFPVFPALITARVHREGATDEEKLRSKMKRYATFARGN PNGNDLSIMPAVNLTRYQYLODLDMISLGRNLKRPVDFWSTTIPLLISADMLHSPLS SUTFLSLQPFPPCAP"

CDS

ORIGIN

Query Match 73.2%; Score 1579.4; DB 9; Length 1962;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

186 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGCTGGAGTGAATTCAGGGCAAGCA 245
222 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGCTGGAGTGAATTCAGGGCAAGCA 281
246 GTCACTGTGTGGGAAGCCCTGCTGGAAGCTGTCTGGAGTGTCCCTTGTGCT 305
282 GTCACTGTGTGGGAAGCCCTGCTGGAAGCTGTCTGGAGTGTCCCTTGTGCT 341
306 CCCCCTGGGATCCCTGCGATTTCAGAACCCGACCTGCAATCCCTGGGATTAATTG 365
342 CCCCCTGGGATCCCTGCGATTTCAGAACCCGACCTGCAATCCCTGGGATTAATTG 401
366 CGAGAGCCACCTCCACCTAATTTGTGCTCCAGAACTCGAGTGTGCTCTTAAAT 425
402 CGAGAGCCACCTCCACCTAATTTGTGCTCCAGAACTCGAGTGTGCTCTTAAAT 461
426 CAACACATGCTCAAGGTGATTAACCGAATTCGAGTGTGAGAGACTGCTTACCG 485
462 CAACACATGCTCAAGGTGATTAACCGAATTCGAGTGTGAGAGACTGCTTACCG 521
486 AACATTTAGCCCTGCGACCCGATACAGGCTCGAAGCTCCCTGTGTGTGTTG 545
522 AACATTTAGCCCTGCGACCCGATACAGGCTCGAAGCTCCCTGTGTGTGTTG 581
546 CCAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAATGGTTCGACCTGGCTGCC 605
582 CCAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAATGGTTCGACCTGGCTGCC 641
606 TATGAGAGCTGCTGTGTGTGCTGCTCAATACCGGCTAGGAATATTTGTTCTTACC 665
642 TATGAGAGCTGCTGTGTGTGCTGCTCAATACCGGCTAGGAATATTTGTTCTTACC 701
666 ACATGGATTCAGCATGCTCCGGGGAACCTGGGCTTCAAGACCAAGTGGCTGTCTGCC 725
702 ACATGGATTCAGCATGCTCCGGGGAACCTGGGCTTCAAGACCAAGTGGCTGTCTGCC 761
726 TGGGTCCAGAAACATGAGTTCTTGGTGGGAGACCCAGACTGTGACCATCTTTGGC 785
762 TGGGTCCAGAAACATGAGTTCTTGGTGGGAGACCCAGACTGTGACCATCTTTGGC 821
786 GAGTCCGGGAGGACATAGTGTCTTAACTTAACTGCTCCCATGGCCAAAGGCTTA 845


```

Db      822 GAGTCGCGGAGGACCAATAGTCTTTCTACTCTTAATCTCTCCCATGCGCAAAAGCTTAA 881
Qy      846 TTTCACAAAGCCATCATAGAGAGTGGGTGGCCATCATCCCTTACTGAGAGCCCATGAT 905
Db      882 TTTCACAAAGCCATCATAGAGAGTGGGTGGCCATCATCCCTTACTGAGAGCCCATGAT 941
Qy      906 TATGAGAGAGTGAAGACCTGAGAGGTGTGACATTTCTGTGTAAACAATGCGTCAAG 965
Db      942 TATGAGAGAGTGAAGACCTGAGAGGTGTGACATTTCTGTGTAAACAATGCGTCAAG 1001
Qy      966 TCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGTGTGACCTGACG 1025
Db      1002 TCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGTGTGACCTGACG 1061
Qy      1026 CAGAAAACAAATCTTTCACTGAGAGTGTGATGTGTCTTTCTTCTTAATAGGCTCTA 1085
Db      1062 CAGAAAACAAATCTTTCACTGAGAGTGTGATGTGTCTTTCTTCTTAATAGGCTCTA 1121
Qy      1086 GATCTATTGTCTCAGAAAGCATTTTAAAGCATTCCTTCATCATGAGAGTCAATAAACAC 1145
Db      1122 GATCTATTGTCTCAGAAAGCATTTTAAAGCATTCCTTCATCATGAGAGTCAATAAACAC 1181
Qy      1146 GAGTGTGGCTTCTGCTGCTCTATGAAGAGGCTCTGAGATCTCAAGTGTCCAAACAG 1205
Db      1182 GAGTGTGGCTTCTGCTGCTCTATGAAGAGGCTCTGAGATCTCAAGTGTCCAAACAG 1241
Qy      1206 TCCCTTGGCCCTCATCTGATACAAACATCTGACATCCCGCTCAATATTTGCACCTT 1265
Db      1242 TCCCTTGGCCCTCATCTGATACAAACATCTGACATCCCGCTCAATATTTGCACCTT 1301
Qy      1266 GTGGCTAATGAATTAATCTTCATGACAAAGCATCCCTGACATGAATCCGAGACGTCTTG 1325
Db      1302 GTGGCTAATGAATTAATCTTCATGACAAAGCATCCCTGACATGAATCCGAGACGTCTTG 1361
Qy      1326 GACTTGTCTGAGATGTGTCTTTGTGTGTCCTGACATGATCAAGCTGATATCAACA 1385
Db      1362 GACTTGTCTGAGATGTGTCTTTGTGTGTCCTGACATGATCAAGCTGATATCAACA 1421
Qy      1386 GATGCTGTGACCTGTCTACTTCTATGAGTTTCGGACCGGCTCAAGTCTTTGAAGAC 1445
Db      1422 ----- 1421
Qy      1446 ACGAAGCGGGCTTTGTCAAAAGCCAGACACGCTGATGAAGTCCGCTTGTTCGGTGT 1505
Db      1422 ----- 1421
Qy      1506 GCCTTCTGAAGGGGACATTTGTATGTTGGAAGAGCCAGGAGAGAGAGATTTACTG 1565
Db      1422 -----GAAAGAGCCACCGAAGAGAGAGATTTACTG 1451
Qy      1566 AGCCGGAAGATGATAAATACTGGGCTACCTTTGCTCGAACCGGAATCTAATGGGAAC 1625
Db      1452 AGCCGGAAGATGATAAATACTGGGCTACCTTTGCTCGAACCGGAATCTAATGGGAAC 1511
Qy      1626 GACCTGTCTCTGTGGCCAGCTTATAATCTGACCTGAGAGTACTCTCAGCTGGAATTTGAAC 1685
Db      1512 GACCTGTCTCTGTGGCCAGCTTATAATCTGACCTGAGAGTACTCTCAGCTGGAATTTGAAC 1571
Qy      1686 ATGAGCCTCGACAGAGTCAAAAGACCGGGGTGATTTTGGACCAACATATCCCC 1745
Db      1572 ATGAGCCTCGACAGAGTCAAAAGACCGGGGTGATTTTGGACCAACATATCCCC 1631
Qy      1746 CTGATCTCTGTCTGCTCGACATGCTCAAGTCTCTTTCTTCTTAACTTTCTCTCT 1805
Db      1632 CTGATCTCTGTCTGCTCGACATGCTCAAGTCTCTTTCTTCTTAACTTTCTCTCT 1691
Qy      1806 CTCCTCCAGCCCTTTCTTTTCTTTGTGCTCTCTGAGAGATCTCTTCTGAGCTCAAGCTGC 1865
Db      1632 CTCCTCCAGCCCTTTCTTTTCTTTGTGCTCTCTGAGAGATCTCTTCTGAGATTTG 1751
Qy      1866 TTTCCCTTCTCTCCCATTAATTTCTCCGCAATCAATTAAGCTTTCTGAGCTCAAGCTGC 1925

```

```

Db      1752 TTTCCCTTCTCTCCCATTAATTTCTCCGCAATCAATTAAGCTTTCTGAGCTCAAGCTGC 1811
Qy      1926 TTTCTATGGGAGATCTCTTGCAAAAGAGCTGCTTGGCTGATATTTATGAGACTTGAAT 1985
Db      1812 TTTCTATGGGAGATCTCTTGCAAAAGAGCTGCTTGGCTGATATTTATGAGACTTGAAT 1871
Qy      1986 GATCCTTACAGAAATCTTTTCAACATCAAAAGTCAATTTGTCTTGAAGGCAACA 2045
Db      1872 GATCCTTACAGAAATCTTTTCAACATCAAAAGTCAATTTGTCTTGAAGGCAACA 1931
Qy      2046 TTTCTTCAATTAATTTGGAAGAGGCTGACC 2076
Db      1932 TTTCTTCAATTAATTTGGAAGAGGCTGACC 1962

```

RESULT 12
BC069548
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

1 (bases 1 to 1896)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marnina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Toshylyuk,S., Carminci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McMan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulik,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,B., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,Y.S., Krzywicki,M.I., Skalska,U., Smalins,D.B., Scherch,A., Schein,J.E., Jones,S.J. and Matra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 1896)
Strausberg,R.
Direct Submission
Submitted (29-Apr-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Baylor Human Genome Sequencing Center
CDNA Library Preparation: Baylor Human Genome Sequencing Center
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIND)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulik, S.W., Loulseged, H., Kowis, C.R., Sneed, A.D., Martin, R.G., Muzny, D.M., Navevati, A.N., Gibbs, R.A.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/INL at: <http://image.lnl.gov>
Series: IRBR Plate: 4 Row: f Column: 1.
Location/Qualifiers
1. 1896

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:97193 IMAGE:7262439"
/issue_type="PCR rescued clones"
/clone_id="NIH MGC 244"
/note="Vector: pPCR-Script Amp SK(+)"
1. 1896

/gene="FLJ31547"
/db_xref="LocusID:221223"
120. 1697
/gene="FLJ31547"

CDS

/codon_start=1
/product="hypotheetical protein FLJ31547"
/protein_id="AAH69548.1"
/db_xref="GI:46854431"
/translation="MSGNVHPGQILIMAVLIAAPTKPSAEGQRRLTGQKQ
VTVLSRPVNVPLGVPAPAPRGLRFTNPQASFWDLKRAISTPMLCLQNSML
LQOHMKNVHPKRVSEDCLYLNTIAPAHADTGSKLPLVWPEGAFTGSAISIDGS
ALAAEDVAVVQYRLGIFFTWDQHPQNMARQDVALSWOKIIEFGDPS
SVTIFESGALISVSLIISPMACKLPHRAIMESGVALIPLYEADHYEKSDIOYVAH
FGNNASDEBALRCLRTKPSKELTISQKTSFRTVVDGAFPEPNEPDLISQKFA
ISIIIVNNHCEGFLPMKEAPEILISGSKSLALHIONLIIPROYLIHIVANEYFD
KSLIETIRSDLDLIGDFPVVPLITATYHREKATBEKLSRQMKMTWATFATGN
PQNDSLIMPAYNLITQYIQLDLNSLQRLKEPRVDFTWTSTPILISASDMLSPLS
SITPLSLDLPFFFCAP"

ORIGIN

Query Match 71.5%; Score 1543.4; DB 9; Length 1896;
Best Local Similarity 91.9%; Pred. No. 0;
Matches 1704; Conservative 0; Mismatches 1; Indels 150; Gaps 1;
186 GGGCCCTTCGCTGAAGGCGCAAGAGAAACACAGGCTGGATTCAGGGCAAGCA 245
192 GGGCCCTTCGCTGAAGGCGCAAGAGAAACACAGGCTGGATTCAGGGCAAGCA 251
246 GTCACTGCTGCTGGAGGCGCTGCTGCTGGAAGGCTTCCTCGAGTCCCTTGGCT 305
252 GTCACTGCTGCTGGAGGCGCTGCTGCTGGAAGGCTTCCTCGAGTCCCTTGGCT 311
306 CCCCCGCTGGAGTCCCTGCTGCTGGAAGGCTTCCTCGAGTCCCTTGGCT 365
312 CCCCCGCTGGAGTCCCTGCTGCTGGAAGGCTTCCTCGAGTCCCTTGGCT 371
366 CGAAGAGCACCTCTACCTTAATTTGCTCCGAACCTCAGAGTGGCTCTTAAGAT 425
372 CGAAGAGCACCTCTACCTTAATTTGCTCCGAACCTCAGAGTGGCTCTTAAGAT 431
426 CAACATGCTCAAGGTGCTATTAACCGAATTCGAGTGTGAGAGACTGCTTACCTG 485
432 CAACATGCTCAAGGTGCTATTAACCGAATTCGAGTGTGAGAGACTGCTTACCTG 491
486 AAATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
492 AAATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
546 CCAAGAGTGGCTTCAGAGCTGCTCAGGCTCCTCAATCTTTGATGGGTCCGCTGGCTGC 605
552 CCAAGAGTGGCTTCAGAGCTGCTCAGGCTCCTCAATCTTTGATGGGTCCGCTGGCTGC 611
606 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
612 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
666 AATAGGATCAGATGCTCCGGGGAATCTGGGCTTCAGAGCAAGAGTGGCTGCTGCTC 725
672 AATAGGATCAGATGCTCCGGGGAATCTGGGCTTCAGAGCAAGAGTGGCTGCTGCTC 731

QY 726 TGGGTCCAGAGAACATGAGTTCTTGGTGGGAGCCCAAGCTGTGACCATCTTTGGC 785
DB 732 TGGGTCCAGAGAACATGAGTTCTTGGTGGGAGCCCAAGCTGTGACCATCTTTGGC 791
QY 786 GAGTCCGGGGGAGCCATAGTGTCTTCTGCTTATCTGTCTCCCATGGCCAAAGCTTGA 845
DB 792 GAGTCCGGGGGAGCCATAGTGTCTTCTGCTTATCTGTCTCCCATGGCCAAAGCTTGA 851
QY 846 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGGAGGCCCATGAT 905
DB 852 TTCCCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGGAGGCCCATGAT 911
QY 906 TATGAGAGAGTGAAGACCTGAGGTGCTGCACATTTCTGTGTAAACATGCTCAGAC 965
DB 912 TATGAGAGAGTGAAGACCTGAGGTGCTGCACATTTCTGTGTAAACATGCTCAGAC 971
QY 966 TCTGAGGCGCTCGAGGTGCTGAGGACAAACCCCTCCAAAGGCTGTGACCTTACG 1025
DB 972 TCTGAGGCGCTCGAGGTGCTGAGGACAAACCCCTCCAAAGGCTGTGACCTTACG 1031
QY 1026 CAGAAAACAAAGTCTTCACTCGAGGTGATGAGTCTTCTTCTTAATGAGCCTTA 1085
DB 1032 CAGAAAACAAAGTCTTCACTCGAGGTGATGAGTCTTCTTCTTAATGAGCCTTA 1091
QY 1086 GATCATTTGCTCGAAGACATTTAAAGCAATCTTCCATCATGAGTCAATTAACAC 1145
DB 1092 GATCATTTGCTCGAAGACATTTAAAGCAATCTTCCATCATGAGTCAATTAACAC 1151
QY 1146 GAGTGTGGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
DB 1152 GAGTGTGGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
QY 1206 TCCCTTCCCTCCATCTGATTAACAAACATCTGACATCCCGCTCAGATTTGCACTT 1265
DB 1212 TCCCTTCCCTCCATCTGATTAACAAACATCTGACATCCCGCTCAGATTTGCACTT 1271
QY 1266 GTGGCTAATGATTAATCTTCCATGACCAAGCACTCCCTGACCTGAATTCGAGACATCTTCTG 1325
DB 1272 GTGGCTAATGATTAATCTTCCATGACCAAGCACTCCCTGACCTGAATTCGAGACATCTTCTG 1331
QY 1326 GACTTGTGAGATGATGTTCTTGTGTGCTCCCTGACCTGATCAGCTGATATCAGAGA 1385
DB 1332 GACTTGTGAGATGATGTTCTTGTGTGCTCCCTGACCTGATCAGCTGATATCAGAGA 1391
QY 1386 GATGCTGTGACACTGTCTATCTTATGAGTTTCGACCCGCTCAGTCTTTGAAGAC 1445
DB 1392 ----- 1391
QY 1446 ACGAAGCCGACTTTTGTCAAAAGCCGACAGCTGATGAAGTCCGCTGTGTGGTGGT 1505
DB 1392 ----- 1391
QY 1506 GCTTCTTAAGAGGAGCACTTGTATGTTGGAAGAGCCAGAGAGAGAGATTACTG 1565
DB 1392 -----GAAGAGCCACGAGAGAGAGAGATTACTG 1421
QY 1566 AGCCGGAAGTATGAATAATCTGGCTACTTTGTGTAACCGGGAATCTTAATGGGAAC 1625
DB 1422 AGCCGGAAGTATGAATAATCTGGCTACTTTGTGTAACCGGGAATCTTAATGGGAAC 1481
QY 1626 GACTGTCTCTGTGGCCAGCTTAATATCTGACTGAGAGTACCTTCAGCTGGAATTGAAC 1685
DB 1482 GACTGTCTCTGTGGCCAGCTTAATATCTGACTGAGAGTACCTTCAGCTGGAATTGAAC 1541
QY 1686 ATGAGCTCTCGAGCAGAGCTCAAAAGAACCGGGGTGGAATTTTGGACCAAGACATCCCC 1745
DB 1542 ATGAGCTCTCGAGCAGAGCTCAAAAGAACCGGGGTGGAATTTTGGACCAAGACATCCCC 1601
QY 1746 CTGATCTGTCTGCTCGAGCATGCTCAGAGTCCCTTCTTCTTAACTTCTCTCT 1805
DB 1602 CTGATCTGTCTGCTCGAGCATGCTCAGAGTCCCTTCTTCTTAACTTCTCTCT 1661
QY 1806 CTCTCTCAGCCTTCTTTTCTTTGTGCTCCTTGAAGAGTTATCTTCTGTGATTTGG 1865

|||||
Db 1662 CTCCTCAGCCCTTCTTTCTTTGCTCCTTGGAAGTATCTTTCTGATTTTG 1721
Qy 1866 TTTCCCTCTCTCCCTAATTTTCCCGCATCATAGCTCTTTCTGAGCTCAGCTGC 1925
Db 1722 TTTCCCTCTCTCTCCCTAATTTTCTCCGCAATCATATAGCTCTTTCTGAGCTCAGCTGC 1781
Qy 1926 TTTCTATGGGAGATCTCTGCAAAACAAGCTGCTTTCGCTGATATTTTATGACTTGAAGAT 1985
Db 1782 TTTCTATGGGAGATCTCTGCAAAACAAGCTGCTTTCGCTGATATTTTATGACTTGAAGAT 1841
Qy 1986 GATCCTTACAGAAATTTCTTTCAACATCAAAAGATGCAATTTGTTGAAAGGCAA 2040
Db 1842 GATCCTTACAGAAATTTCTTTCAACATCAAAAGATGCAATTTGTTGAAAGGCAA 1896

RESULT 13
CQ733044 1716 bp DNA linear PAT 03-FRB-2004
LOCUS Sequence 18978 from Patent WO02068579.
DEFINITION CQ733044
ACCESSION CQ733044 GI:42314969
VERSION CQ733044.1 GI:42314969
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
cheerof
JOURNAL Patent: WO 02068579-A 18978 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Source 1..1716
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 66.0%; Score 1425; DB 6; Length 1716;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 179 AAAAACTGGGCTTGTGCTGAAGGCGCACAGAGAACACAGGCTGGAGTTCAGGG 238
Db 21 AAACAAAGGGCTTGTGCTGAAGGCGCACAGAGAACACAGGCTGGAGTTCAGGG 80
Qy 239 CAAGCAAGTCACTGCTGGAGAGCGCTGTGCTGGAACGCTTCTCCGAGTCCCTT 298
Db 81 CAAGCAAGTCACTGCTGGAGAGCGCTGTGCTGGAACGCTTCTCCGAGTCCCTT 140
Qy 299 TGTCTGCTCCCGCTGGAGTCCCTGCAATTTACGAACCGCAGCTGCAATCCCTGGGA 358
Db 141 TGTCTGCTCCCGCTGGAGTCCCTGCAATTTACGAACCGCAGCTGCAATCCCTGGGA 200
Qy 359 TAACTTGGAGAGGCACTCTCACTCAATTTGTGCTTCCAGAACTCAGAGTGCCT 418
Db 201 TAACTTGGAGAGGCACTCTCACTCAATTTGTGCTTCCAGAACTCAGAGTGCCT 260
Qy 419 CTTAGATCAACAGCTCAAGGTCATTAACCGAAATTCGAGTGTCAAGAGACTGCCT 478
Db 261 CTTAGATCAACAGCTCAAGGTCATTAACCGAAATTCGAGTGTCAAGAGACTGCCT 320
Qy 479 CTACCTGAACATCTATGGGCTGCGCAGCGCATACAGGCTCAAGCTCCCGTCTGGT 538
Db 321 CTACCTGAACATCTATGGGCTGCGCAGCGCATACAGGCTCAAGCTCCCGTCTGGT 380
Qy 539 GTGGTTCCAGAGAGTGCCTTCAAGCTGAGCTCAAGCTTCAATCTTTGATGGTCCGCCCT 598
Db 381 GTGGTTCCAGAGAGTGCCTTCAAGCTGAGCTCAAGCTTCAATCTTTGATGGTCCGCCCT 440

Qy 599 GGCTGCTATGAGAGCTGCTGCTGTGTGCTGCTCAAGTACCGGCTAGGAATATTTGTT 658
Db 441 GGCTGCTATGAGAGCTGCTGCTGTGTGCTGCTCAAGTACCGGCTAGGAATATTTGTT 500
Qy 659 CTTTACACATGGGATTCAGCATGCTCCGGGAACTGGGCTTCAAGAGACAGGTGCTGC 718
Db 501 CTTTACACATGGGATTCAGCATGCTCCGGGAACTGGGCTTCAAGAGACAGGTGCTGC 560
Qy 719 TCTGTCTGGGCTGCAAGGAACATGAGTCTTGGTGGGAGCCAGCTGTGACCAT 778
Db 561 TCTGTCTGGGCTGCAAGGAACATGAGTCTTGGTGGGAGCCAGCTGTGACCAT 620
Qy 779 CTTTGGAGAGTCCGGGAGCCATTAAGTGTTCATGCTTATATCTGCTCCATGCGCAA 838
Db 621 CTTTGGAGAGTCCGGGAGCCATTAAGTGTTCATGCTTATATCTGCTCCATGCGCAA 680
Qy 839 AGGCTTATTTCCAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTACCTGAGGC 898
Db 681 AGGCTTATTTCCAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTACCTGAGGC 740
Qy 899 CCATGATTTATGAGAGAGTGAAGGAGCTGCAAGTGGTTGCAATTTCTGTGTGAAGATGC 958
Db 741 CCATGATTTATGAGAGAGTGAAGGAGCTGCAAGTGGTTGCAATTTCTGTGTGAAGATGC 800
Qy 959 GTCAAGCTCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTTCAAGAGAGCTGTGAC 1018
Db 801 GTCAAGCTCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCTTCAAGAGAGCTGTGAC 860
Qy 1019 CCTCAGCCGAGAAACAAAGTCTTTCACTGAGGTGTGATGCTTCTTCTTATGA 1078
Db 861 CCTCAGCCGAGAAACAAAGTCTTTCACTGAGGTGTGATGCTTCTTCTTATGA 920
Qy 1079 GCCTTATATCTATGCTGCTGAGAAAGCAATTTAAACAAATTCCTTCAATCATCGAGTCAA 1138
Db 921 GCCTTATATCTATGCTGCTGAGAAAGCAATTTAAACAAATTCCTTCAATCATCGAGTCAA 980
Qy 1139 TAAACAGAGTGTGCTTCTGCTGCTCATGAGAGAGGCTCCTGAGATCTTCAGTGTCTC 1198
Db 981 TAAACAGAGTGTGCTTCTGCTGCTCATGAGAGAGGCTCCTGAGATCTTCAGTGTCTC 1040
Qy 1199 CAACAAATCCTTGTGCTTCACTGATTAACAAACATCTGCAACATCCGCTCAATATT 1258
Db 1041 CAACAAATCCTTGTGCTTCACTGATTAACAAACATCTGCAACATCCGCTCAATATT 1100
Qy 1259 GCACTTGTGCTATGAAATCTTCCATGACAAAGACCTCCGATGAAATCCGAGACAG 1318
Db 1101 GCACTTGTGCTATGAAATCTTCCATGACAAAGACCTCCGATGAAATCCGAGACAG 1160
Qy 1319 TCTTCTGACCTTGTGAGATGTGTCTTGTGTGCTGCTGACATGATCAAGCTGATTA 1378
Db 1161 TCTTCTGACCTTGTGAGATGTGTCTTGTGTGCTGCTGACATGATCAAGCTGATTA 1220
Qy 1379 TCACAGAGATGTGTGACCTGTCTACTTCTATAGATTTTGGACCGGCTCAATGCTT 1438
Db 1221 TCACAGAGATGTGTGACCTGTCTACTTCTATAGATTTTGGACCGGCTCAATGCTT 1280
Qy 1439 TGAAGACACAGAGCGGCTTTTGTAAAGCCGACCAAGCTGATGAAGTCCGTTGTGT 1498
Db 1281 TGAAGACACAGAGCGGCTTTTGTAAAGCCGACCAAGCTGATGAAGTCCGTTGTGT 1340
Qy 1499 CGGTGTGCTCTTCTGAAAGGGGAGCAATGTTATGTTTGAAGAGGACGAGAGAGAGNA 1558
Db 1341 CGGTGTGCTCTTCTGAAAGGGGAGCAATGTTATGTTTGAAGAGGACGAGAGAGAGNA 1400
Qy 1559 GTTACTGAGCGGAGATGAATTAATCTGGCTTACCTTTGCTGAAACGGGA 1611
Db 1401 GTTACTGAGCGGAGATGAATTAATCTGGCTTACCTTTGCTGAAACGGGA 1453

RESULT 14
AX958392 1857 bp DNA linear PAT 14-JAN-2004
LOCUS AX958392
DEFINITION Sequence 20 from Patent WO0226988.

TITLE Direct Submission
JOURNAL Submitted (04-AUG-2004) Masao Miyazaki, RIKEN, Frontier System
Research, Supra-Biomolecular System Research Group, Hirosewa 2-1,
Wako-shi, Saitama 351-0198, Japan
(E-mail: mmiyazaki@postman.riken.jp, Tel: 81-48-467-9619,
Fax: 81-48-467-9620)

FEATURES Location/Qualifiers

Source

1..1728

/organism="Canis familiaris"

/mol_type="mrna"

/db_xref="taxon:9615"

/chromosome="2"

1..1728

/gene="cauxin"

1..1728

/gene="cauxin"

/product="carboxylesterase-like urinary excreted protein"

/protein_id="BAD35015.1"

/db_xref="GI:51014275"

/translation="MSGENGHIGQTLINAVWTLAATSPADAPVRSRLGWKQ

ATVLSSTFVNFLGIPEAPPLGPKRFRKPKALMDSRATSPKLCLONSVWL

SDQHLKVHPMLVSESDCLYNTIYAPAHNTGSKLPVWVPPGAFETGSAIFDGS

ALAAVEDVLIYTOYRLGIFGFKTGDHPGNWAFIDQALATVQENIEFGDPH

SVTIFGESGAISSGLVSPWASGLFKHAIWESGVAIIPFLAPDEBNDLOIYAR

ICGNVSDSVALLQCLRAKSSRELDINKTKSPTRVYDPPFPDPLDLEKTPNS

IPSVGVNNHECGFLIPKPEPPEITIGSKSLALHLIRVLIHPQYLYLVADQYFN

KHSPEIRDSFLDLGVFVPPGVVTAHYHDAQAPVYFVFOHPBELKQTRAPV

KADHDEIRFVPGAFILKNIIVMEGATIEEELSLRKMYRANARPTDPEGLPL

MPAYSOBOYKLDLNIISVQKLKEQVEVFMSDTLPLIMSMSTARPPVPILISVL

LPLFSSAP"

ORIGIN

Query Match 55.5%; Score 1198.4; DB 4; Length 1728;
Best Local Similarity 82.7%; Pred. No. 0;
Matches 1370; Conservative 0; Mismatches 286; Indels 0; Gaps 0;

186 GGDCTTCTGCTGAAGGCGCACAGAGGAACACAGCTGGATTCGATTGAGGCGAAGAA 245
73 GGTTCAGCTGCTGAAGCAACGGTGAAGAGCACAGCTGGATTCGAGGCGAAGCA 132
246 GTCACTGCTGCTGAAGAGCCCTGCTGCTGAAGAGTTCCTGCAAGTCCCTTTGCTGCT 305
133 GCCACTGCTGCTGAAGAGCAACAGCTGCTGCAAGTTCCTGCAAGTCCCTTTGCTGCT 192
306 CCCCCTGCTGAAGTCCCTGCTGAAGTTCGAAACCGGAGCTGCAAGTCCCTTTGCTGCT 365
193 CCCCCTGCTGAAGTCCCTGCTGAAGTTCGAAACCGGAGCTGCAAGTCCCTTTGCTGCT 252
366 CGAGAAGCACTCTCAACCTTAATTTGCTGCTGCAAGTTCGAAAGTGGCTGCTTAAGT 425
253 CGAGATGCCACCTCTATCTTAATTTGCTGCTGCAAGTTCGAAAGTGGCTGCTTAAGT 312
426 CAACATGCTCAAGGTGCAATACCGGAATTCGAGTGTCAAGAGTGGCTTCACTG 485
313 CAACATTTCTCAAGGTGCAATACCGGAATTTGAAGTGTCCGAAGTGGCTTCACTG 372
486 AACATCTATGCTGCTGCCAGCGCCGATTAACAGCTCCAGCTCCCGTCTTGTGTGCTT 545
373 AACATCTATGCTGCTGCCAGCGCCGATTAACAGCTCCAGCTCCCGTCTTGTGTGCTT 432
546 CAGAGAGTGTCTTCAAGAGTGTCTCAAGCTGCTCAATTTTGAAGTGGCTCCGCTGCTG 605
433 CTGAGAGTGTCTTCAAGAGTGTCTCAAGCTGCTCAATTTTGAAGTGGCTCCGCTGCTG 492
606 TATGAGAGTGTCTGCTGCTGCTGCTCAAGTGTCTCAAGTGTCTGAGAAATTTGTTTCTT 665
493 TATGAGAGTGTCTGCTGCTGCTGCTCAAGTGTCTCAAGTGTCTGAGAAATTTTGTGCT 552
666 ACATGGATCAAGTGTCTGCTGCTGCTGCTCAAGTGTCTGAGAAATTTGTTTCTTGTCT 725
553 ACAGGGAGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 612

726 TGGTCCAGAGAAATCATGAGTCTTCCGTTGGGAGACCCAGCTGTGATCACTTTGGC 785
613 TGGTCCAGAGAAATCATGAGTCTTCCGTTGGGAGACCCAGCTGTGATCACTTTGGC 672
786 GAGTCCGCGGAGCAATAGTGTCTGAGTCTTATCTGTCTTCCAGTGGCCAAAGCTTA 845
673 GAGTCCAGAGAAATCATGAGTCTTCCGTTGGGAGACCCAGCTGTGATCACTTTGGC 732
846 TTCCCAAAAGCCATCATGAGAGTGGGAGCAATCCCTTCACTGAGAGCCCATGAT 905
733 TTCCCAAAAGCCATCATGAGAGTGGGAGCAATCCCTTCACTGAGAGCCCATGAT 792
906 TATGAGAGAGTGAAGAGTCAAGTGTGTCACATTTCTGTGTAAACAATCGTCAGAC 965
793 GATGAGAGAGTGAAGAGTGTGTCACATTTCTGTGTAAACAATCGTCAGAC 852
966 TCTGAGGCTCTGAGAGTGTGTCAGACCAAAACCTTCAAGAGCTGTGACCTCAGC 1025
853 TCCGTAGCTCTGAGAGTGTGTCAGAGGCAAAATCCTCGAGAGTGTGTCAGATCAAC 912
1026 CAGAAACCAAGTCTTCACTGAGTGTGTCAGAGTGTGTCAGAGTGTGTCAGAGTCTTA 1085
913 AAGAAACCAAGTCTTCACTGAGTGTGTCAGAGTGTGTCAGAGTGTGTCAGAGTCTTA 972
1086 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATGAGAGTCAATAACAC 1145
973 GACCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATGAGAGTCAATAACAC 1032
1146 GAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205
1033 GAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
1206 TCCCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265
1093 TCCCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
1266 GTGGCTAATGATCTTCACTGATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1325
1153 GTGGCTAATGATCTTCACTGATGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1212
1326 GACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1385
1213 GACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
1386 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1445
1273 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332
1446 AGGAAGCGGCTTTGTCAAGCGAGCAACGCTGATGAGAGTCCGCTTGTGTGCTGCT 1505
1333 AGGAAGCGGCTTTGTCAAGCGAGCAACGCTGATGAGAGTCCGCTTGTGTGCTGCT 1392
1506 GCTTCTGAGAGGAGGAGCAATTTGATGATGAGAGGAGCAACGAGAGAGAGAGTACTG 1565
1393 GCTTCTGAGAGGAGGAGCAATTTGATGATGAGAGGAGCAACGAGAGAGAGAGTACTG 1452
1566 AGCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1625
1453 AGCGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
1626 GACCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1685
1513 GACCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1572
1686 ATGAGCTCTGAGCAGAGACTCAAGAGCCGCGGTGAGATTTTGGACAGCAATCCCC 1745
1573 ATGAGCTCTGAGCAGAGACTCAAGAGCCGCGGTGAGATTTTGGACAGCAATCCCC 1632
1746 CTGATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1805
1633 CTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1692
1806 CTCTCTGAGCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1841

Thu Jun 16 11:02:00 2005

us-10-023-515-1.rge

Page 20

Db 1693 GTGCTCTGCCTTCTATCTCTCCTCTGCTCCTGA 1728

Search completed: June 14, 2005, 22:52:42
Job time : 9387.39 secs

PT surrogate markers, in tissue typing and chromosome mapping.
XX
PS Claim 1; Page 109-110; 122pp; English.
XX

CC The present nucleotide sequence, the coding region of which is also
CC claimed, is that of cDNA encoding human 53010, a novel member of the
CC carboxylesterase family. The invention provides 53010 nucleic acids,
CC antisense molecules, expression vectors, host cells, transgenic animals,
CC 53010 proteins, fusion proteins, antigenic peptides, anti-53010
CC antibodies and methods for detecting the presence of 53010 polypeptides
CC or nucleic acids, of identifying a compound that binds to the 53010
CC polypeptide, and of modulating the activity of the polypeptide. The 53010
CC nucleic acids and polypeptides can act as novel diagnostic and
CC therapeutic agents for controlling disorders involving aberrant or
CC deficient hydrolysis of carboxylic esters. As 53010 mRNA is highly
CC expressed in the central and peripheral nervous system, and its
CC expression is regulated in some rodent pain models, 53010 molecules can
CC also act as novel diagnostic targets and therapeutic agents for
CC controlling neurological disorders, such as pain-related disorders. A
CC claimed method of treating or preventing a disorder (especially a pain-
CC related disorder) characterized by aberrant activity of a 53010-
CC expressing cell involves administering a compound that modulates 53010
CC activity or expression. 53010 nucleic acids are also useful in chromosome
CC mapping, tissue typing, in forensic biology, prognostic assays, in
CC arrays, for detection of variations or mutations, as surrogate markers
CC and in pharmacogenomics
XX
SQ Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;

Query Match 100.0%; Score 2158; DB 6; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACGCGTCGGAACACAGGCTGAGAGCAATGAGAGTAACCAATGTAATTAACCA 60
DB 1 CCACGCGTCGGAACACAGGCTGAGAGCAATGAGAGTAACCAATGTAATTAACCA 60
QY 61 TTTAAAAATTATTCTTATATAAAGTAAGATCCCAATGCCACAGGAGCTTACTTCATCTG 120
DB 61 TTTAAAAATTATTCTTATATAAAGTAAGATCCCAATGCCACAGGAGCTTACTTCATCTG 120
QY 121 CTTCACAATGAGTGTCTTTTCTGATATCTCAAGCCCTCTGTTGGACAAGAGTGGGAA 180
DB 121 CTTCACAATGAGTGTCTTTTCTGATATCTCAAGCCCTCTGTTGGACAAGAGTGGGAA 180
QY 181 AAACCTGGGCTTCTGCTGAAGGGCCAAGAGAAACACAGGCTGGAGTGAATTCAGGGCA 240
DB 181 AAACCTGGGCTTCTGCTGAAGGGCCAAGAGAAACACAGGCTGGAGTGAATTCAGGGCA 240
QY 241 AACCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCGGAGTCCCTTTTG 300
DB 241 AACCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCGGAGTCCCTTTTG 300
QY 301 CTGCTCCCGCGTGGAGATCCTGCGATTTAACGAAACCCGAGCCTGCAATCCGCTGGGATA 360
DB 301 CTGCTCCCGCGTGGAGATCCTGCGATTTAACGAAACCCGAGCCTGCAATCCGCTGGGATA 360
QY 361 ACTTGCGAGAACCACTCTCTACCTTAATTGTGCTCAAGAACTCAAGTGGGCTGCTCT 420
DB 361 ACTTGCGAGAACCACTCTCTACCTTAATTGTGCTCAAGAACTCAAGTGGGCTGCTCT 420
QY 421 TAGATCAACACATGCTCAAGGTGATTAACCGAAATTGGAGTGTCAAGAACTGCTCT 480
DB 421 TAGATCAACACATGCTCAAGGTGATTAACCGAAATTGGAGTGTCAAGAACTGCTCT 480
QY 481 ACCTGAACATCTATGCGGCTGCCACGCGGATACAGGCTCCAAAGCTCCCGCTTGGGT 540
DB 481 ACCTGAACATCTATGCGGCTGCCACGCGGATACAGGCTCCAAAGCTCCCGCTTGGGT 540
QY 541 GGTTCGCCAGAGGTGCTTCAAGACTGGCTCAGCTCCATCTTGAATGGTCCGCCCTGG 600
DB 541 GGTTCGCCAGAGGTGCTTCAAGACTGGCTCAGCTCCATCTTGAATGGTCCGCCCTGG 600

QY 601 CTGCTATAGAGACGTGCTGTGTTGTTGTCAGATCCGGCTAGGAAATTTGGTTCT 660
DB 601 CTGCTATAGAGACGTGCTGTGTTGTTGTCAGATCCGGCTAGGAAATTTGGTTCT 660
QY 661 TCACCAATGAGATCAGATGCTCCGGGAACTGGGCTTCAAGAACAGAGTGGCTGCTC 720
DB 661 TCACCAATGAGATCAGATGCTCCGGGAACTGGGCTTCAAGAACAGAGTGGCTGCTC 720
QY 721 TGTCCCTGGGTCCAGAAAGATCGAGTTCTTGGTGGGGACCCAGCTCTGTACATCT 780
DB 721 TGTCCCTGGGTCCAGAAAGATCGAGTTCTTGGTGGGGACCCAGCTCTGTACATCT 780
QY 781 TTGGCGAGTCCGCGGAGCCATTAAGTGTCTTACTTAAGTCTCCATGGCCAAAG 840
DB 781 TTGGCGAGTCCGCGGAGCCATTAAGTGTCTTACTTAAGTCTCCATGGCCAAAG 840
QY 841 GCTTATTCACAAACCCATCATAGAGATGGGTGGCCATCATCCCTTAAGAGGCC 900
DB 841 GCTTATTCACAAACCCATCATAGAGATGGGTGGCCATCATCCCTTAAGAGGCC 900
QY 901 ATGATTAAGAGAGATGAGACCTGACAGTGTGTCATCTTGTGTGTAACATGCT 960
DB 901 ATGATTAAGAGAGATGAGACCTGACAGTGTGTCATCTTGTGTGTAACATGCT 960
QY 961 CAGACTTGAAGGCTCTGCTGAGGTGCTGAGACAAAACCTCCAAAGAGCTGTGACC 1020
DB 961 CAGACTTGAAGGCTCTGCTGAGGTGCTGAGACAAAACCTCCAAAGAGCTGTGACC 1020
QY 1021 TCAGCCAGAAAACAAAGCTTTCATCTGAGTGTGTAAGTGGCTTTCTTCTTAAGC 1080
DB 1021 TCAGCCAGAAAACAAAGCTTTCATCTGAGTGTGTAAGTGGCTTTCTTCTTAAGC 1080
QY 1081 CTCTAGATCTATGTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGTCAATA 1140
DB 1081 CTCTAGATCTATGTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGTCAATA 1140
QY 1141 AACAGAGTGTGGTCTCTGCTGCTATGAAGAGGCTCTGAGATCTCATGAGTCTCCA 1200
DB 1141 AACAGAGTGTGGTCTCTGCTGCTATGAAGAGGCTCTGAGATCTCATGAGTCTCCA 1200
QY 1201 ACAAGTCCCTTCCCTCATCTGATTAACAAACATCTCGACATCCGCTCAGATTTGCG 1260
DB 1201 ACAAGTCCCTTCCCTCATCTGATTAACAAACATCTCGACATCCGCTCAGATTTGCG 1260
QY 1261 ACCTTGTGCTAATGAATTACTTCATGACAAAGCATCTCCCTGACTGAATCCGAGACGTC 1320
DB 1261 ACCTTGTGCTAATGAATTACTTCATGACAAAGCATCTCCCTGACTGAATCCGAGACGTC 1320
QY 1321 TTCTGACTTGTGAGATGTGTTCTTGTGGCCCTGACATGATCAAGCTCGATATC 1380
DB 1321 TTCTGACTTGTGAGATGTGTTCTTGTGGCCCTGACATGATCAAGCTCGATATC 1380
QY 1381 ACAGAGTGTGCTGCTGACCTGTCTTACTTCTTAGAGTTTGGGCAACGAGCTCAAGTCTTTG 1440
DB 1381 ACAGAGTGTGCTGCTGACCTGTCTTACTTCTTAGAGTTTGGGCAACGAGCTCAAGTCTTTG 1440
QY 1441 AAGACAGAAAGCCGCTTTTGTCAAGCCGACCAACCTGATGAAGTCCGCTTTGTG 1500
DB 1441 AAGACAGAAAGCCGCTTTTGTCAAGCCGACCAACCTGATGAAGTCCGCTTTGTG 1500
QY 1501 GTGGTGTCTTCTGAAGGGGACATTTGTTTCGAAGAGGACGAGAGAGAGAAAT 1560
DB 1501 GTGGTGTCTTCTGAAGGGGACATTTGTTTCGAAGAGGACGAGAGAGAGAAAT 1560
QY 1561 TACTGAGCCGGAAGATGAAATATGAGGCTACCTTTGCTGAAACCGGAAATCTTAATG 1620
DB 1561 TACTGAGCCGGAAGATGAAATATGAGGCTACCTTTGCTGAAACCGGAAATCTTAATG 1620
QY 1621 GGAACGACTGTCTCTGTGGCCAGCTTATATCTGATGAGACGTAACCTCCAGCTGGACT 1680
DB 1621 GGAACGACTGTCTCTGTGGCCAGCTTATATCTGATGAGACGTAACCTCCAGCTGGACT 1680
QY 1681 TGAACATGAGCTTCGACAGAGACTCAAGAAACCGCGGGTGAATTTTGTGACCAACCA 1740

QY	836	CAAAGGCTTATTCACAAGACCATATGAGAGTGGGGTGGCATCATCCCTTACCTGGA	895
Dp	923	CAAAAGCTTATTCACAAGACCATATGAGAGTGGGGTGGCATCATCCCTTACCTGGA	982
QY	896	GGCCCATGATTATGAAAGAGTGAAGACCTGCAAGGTGGTTCACATTTCTGTGTTACAA	955
Dp	983	GGCCCATGATTATGAAAGAGTGAAGACCTGCAAGGTGGTTCACATTTCTGTGTTACAA	1042
QY	956	TGGCTCAGACTCTGAGGGCCCTGCGAGGTGCTGAGGACAAAAACCTTCAGAGAGCTCT	1015
Dp	1043	TGGCTCAGACTCTGAGGGCCCTGCGAGGTGCTGAGGACAAAAACCTTCAGAGAGCTCT	1102
QY	1016	GACCCTCAGCCAGAAACAAAGCTCTTTCACCTGAGTGGTGTGCTTTCTTTCCTTAA	1075
Dp	1103	GACCCTCAGCCAGAAACAAAGCTCTTTCACCTGAGTGGTGTGCTTTCTTTCCTTAA	1162
QY	1076	TGAGCCTCTAATCTATTTGTCTCAGAAAGCATTTAAAGCATTTCTTCATCATCGAGT	1135
Dp	1163	TGAGCCTCTAATCTATTTGTCTCAGAAAGCATTTAAAGCATTTCTTCATCATCGAGT	1222
QY	1136	CAATPAACACGAGTGTGCTTCCTGCTCCATGAAAGAGGCTCCTGAGTCCTCAGTGG	1195
Dp	1223	CAATPAACACGAGTGTGCTTCCTGCTCCATGAAAGAGGCTCCTGAGTCCTCAGTGG	1282
QY	1196	CTCCAAACAAGTCCCTTGGCCTTCATCTGATACAAAACATCCTGCAATCCCGCTCAGTA	1255
Dp	1283	CTCCAAACAAGTCCCTTGGCCTTCATCTGATACAAAACATCCTGCAATCCCGCTCAGTA	1342
QY	1256	TTTTCGACCTTGTCGTATATGAATCTTCACAGACAAAGACCTCCCTGATCTGAATCCGGA	1315
Dp	1343	TTTTCGACCTTGTCGTATATGAATCTTCACAGACAAAGACCTCCCTGATCTGAATCCGGA	1402
QY	1316	CAGTCTTCTGACCTTGACCTTGAGAGATGTGTTCTTTGTGGTCCCTGACATGATCAGACTCG	1375
Dp	1403	CAGTCTTCTGACCTTGACCTTGAGAGATGTGTTCTTTGTGGTCCCTGACATGATCAGACTCG	1462
QY	1376	ATATCAACAGAGATGTGTGTGACCTGTCTACTTCTATGATGTTTGGCACCCGGCTCAGTG	1435
Dp	1463	ATATCAACAGAGATGTGTGTGACCTGTCTACTTCTATGATGTTTGGCACCCGGCTCAGTG	1522
QY	1436	CTTTGAGAGACACGAAGCCGGCTTTTGTCAAAAGCCGACACAGCTGATGAAGTCCGCTTTGT	1495
Dp	1523	CTTTGAGAGACACGAAGCCGGCTTTTGTCAAAAGCCGACACAGCTGATGAAGTCCGCTTTGT	1582
QY	1496	GTTTCGGTGTGTCCTTCTCTGAAGGGGGACATTTGTTATGTTTCGAGAGAGCCACGAGAGAGA	1555
Dp	1583	GTTTCGGTGTGTCCTTCTCTGAAGGGGGACATTTGTTATGTTTCGAGAGAGCCACGAGAGAGA	1642
QY	1556	GAACTTACTGAGCCGGAGAAATGAATTAATGCGGCTACCTTTTCTCGAACCCGGGAATCC	1615
Dp	1643	GAACTTACTGAGCCGGAGAAATGAATTAATGCGGCTACCTTTTCTCGAACCCGGGAATCC	1702
QY	1616	TAAATGAGAACAGACCTGTCTCTGTGGCCAGCTTATTAATCTGACGAGTACCTCGAGCT	1675
Dp	1703	TAAATGAGAACAGACCTGTCTCTGTGGCCAGCTTATTAATCTGACGAGTACCTCGAGCT	1762
QY	1676	GGAATTGAACATGAGCTTCGAGACAGAGACTCAAAAGAACCCGCGGTGATTTTGGACGAG	1735
Dp	1763	GGAATTGAACATGAGCTTCGAGACAGAGACTCAAAAGAACCCGCGGTGATTTTGGACGAG	1822
QY	1736	CACCATCCCCCTGATNCTGTCTGTGCTCGGACATGCTCAGAGTCTCTTTTCTTCTTAAC	1795
Dp	1823	CACCATCCCCCTGATNCTGTCTGTGCTCGGACATGCTCAGAGTCTCTTTTCTTCTTAAC	1882
QY	1796	TTTTCCTCTCTCTCCGACGCTTTCTTTTTCCTTGTGCTCCTTGAAGAGTTATCTTCT	1855
Dp	1883	TTTTCCTCTCTCTCCGACGCTTTCTTTTTCCTTGTGCTCCTTGAAGAGTTATCTTCT	1942
QY	1856	GTCATTTTGTGTTCCCTCTCTCTCCCAATTTTCTCCCGCAATCATTAAGCTTCTTCTGGA	1915
Dp	1943	GTCATTTTGTGTTCCCTCTCTCTCCCAATTTTCTCCCGCAATCATTAAGCTTCTTCTGGA	2002

Oy		1916	GCTAGCGTGTTCATTAGGGGATCCTTGCAAAAACAGCTGCTTGCATATTTATNG	1975
Dd		2003	GCTAAGCTGCTTTCTTAGGGGATCCTTGCAAAACAAGCTGCTTGCATATTTATNG	2062
Oy		1976	ACTTAGGAATGATCCCTTACAGAATTCCTTTCAACATCAAAAAAGTCAATTGCTTGAA	2035
Dd		2063	ACTTAGGAATGATCCCTTACAGAATTCCTTTCAACATCAAAAAAGTCAATTGCTTGAA	2122
Oy		2036	GGCAACAAGATTTCTTAATAATTGGAAGAGGCGTGGCTATTAGTGTCAATAAT	2095
Dd		2123	GGCAACAAGATTTCTTAATAATTGGAAGAGGCGTGGCTATTAGTGTCAATAAT	2182
Oy		2096	GGTTTGTAACTCATATGAATAATAAACAGATATGTAATAATAGGAAAA	2142
Dd		2183	GGTTTGTAACTCATATGAATAATAAACAGATATGTAATAATAGGATA	2229
<hr/>				
RESULT_5				
ADBB62095				
ID	ADBB62095	standard; cDNA; 2092 BP.		
XX				
AC	ADB62095;			
XX				
DT	04-DEC-2003	(first entry)		
XX				
DB	Human cDNA encoding cIone BRAWH20021910.			
XX				
KW	Human; ss; gene; pharmaceutical; diagnostic; gene therapy;			
KW	tissue regeneration; cell regeneration; membrane protein;			
KW	signal transduction-related protein; transcription-related protein;			
KW	osteoporosis; neurological disease; cancer; tumour.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	CDS	394..1803		
FT		/tag= a		
XX		/product= "CIone BRAWH20021910 protein"		
EP	EP1308459-A2.			
PB	07-MAY-2003.			
XX				
PF	28-MAR-2002; 2002EP-00007401.			
XX				
PR	05-NOV-2001; 2001JP-00379298.			
PR	25-JAN-2002; 2002US-00350978.			
PA	(HELI-) HELIX RES INST.			
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.			
PI	Iomgai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,			
PI	Yamamoto J, Isono Y, Hiro Y, Otuka K, Nagai K, Irie R, Tamechika I,			
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;			
DR	WPI; 2003-450961/43.			
DR	P-PDSB; ADB64065.			
XX				
PT	New polynucleotides and polypeptides, useful for developing a diagnostic			
PT	marker or medicines for regulation of their expression and activity, or			
PS	as targets of gene therapy.			
XX				
PS	Claim 1; Page; 222pp; English.			
XX				
<hr/>				
CC	The invention discloses a polynucleotide comprising a sequence selected			
CC	from 1970 fully defined nucleotide sequences which encode novel			
CC	polypeptides. Also claimed is a polypeptide encoded by the polynucleotide			
CC	or its partial peptide, an antibody binding to the polypeptide or peptide			
CC	of the polynucleotide, immunologically assaying the polypeptide or			
CC	peptide of the polynucleotide by contacting the polypeptide or peptide			
CC	with the antibody of the encoded protein, and observing the binding			
CC	between the two, a transformant carrying the polynucleotide in an			
CC	expressible manner and an antisense polynucleotide. The oligonucleotide			

is useful as a primer for synthesizing the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotides and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a cDNA of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 2092 BF; 464 A; 552 C; 497 G; 579 T; 0 U; 0 Other;

Query Match 88.8%; Score 1917; DB 10; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1940; Conservative 0; Mismatches 5; Indels 10; Gaps 1;

```
QY 186 GGGCCCTTCTGCTGAAGGGCCACAGAGAAACACCGCTGGAGATTCAGGCGAAGCAA 245
DB 138 GGGCCCTTCTGCTGAAGGGCCACAGAGAAACACCGCTGGAGATTCAGGCGAAGCAA 197
QY 246 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAAGTGTCTCTGGAGTCCCTTTGTGCT 305
DB 198 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAAGTGTCTCTGGAGTCCCTTTGTGCT 257
QY 306 CCCCCCTGGAGATCCCTGCGATTTAGGAACCCGACCTGCATGCGCTTGGGATTAATTG 365
DB 258 CCCCCCTGGAGATCCCTGCGATTTAGGAACCCGACCTGCATGCGCTTGGGATTAATTG 317
QY 366 CGAAGACCACTCTACCCCTAATTT-----GTGCTCTCAAGAACTCGAAGTGGT 415
DB 318 CGAAGACCACTCTACCCCTAATTTGTGAAGACAGATGCTCGAAGTCTCGAAGTGGT 377
QY 416 GCTTTAGATCAACACATGCTCAAGGTGATTAACCGGAATTCGAGTGCAGAAAGCTG 475
DB 378 GCTTTAGATCAACACATGCTCAAGGTGATTAACCGGAATTCGAGTGCAGAAAGCTG 437
QY 476 CCTCTACCGAATCATATAGCGCTGCGCAAGCCGATACAGGCTCCAGCTCCCGTCTT 535
DB 438 CCTCTACCGAATCATATAGCGCTGCGCAAGCCGATACAGGCTCCAGCTCCCGTCTT 497
QY 536 GGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGAGCTCACTCTTGTATGGGTCCGC 595
DB 498 GGTGTGTGTTCCAGAGAGTGCCTTCAAGACTGAGCTCACTCTTGTATGGGTCCGC 557
QY 596 CCTGGCTGCTATAGGAAGTGTGTTGTGCTGCTCAAGTACCGGCTAAGAAATTTTGG 655
DB 558 CCTGGCTGCTATAGGAAGTGTGTTGTGCTGCTCAAGTACCGGCTAAGAAATTTTGG 617
QY 656 TTTCTTCAACCAATGAGGATCAGATGCTCCGGGGAACCTGGGCTTCAAGAACCAAGTGG 715
DB 618 TTTCTTCAACCAATGAGGATCAGATGCTCCGGGGAACCTGGGCTTCAAGAACCAAGTGG 677
QY 716 TGTCTGTGCTGAGTCCAGAAACATGAGTTCCTCGTGGGAGCCCAAGCTCTGTAGC 775
DB 678 TGTCTGTGCTGAGTCCAGAAACATGAGTTCCTCGTGGGAGCCCAAGCTCTGTAGC 737
QY 776 CATCTTTGGCGAGTCCGGGAGCCATAAGTGTTCCTAGCTTATATCTGTCTCCATGGC 835
DB 738 CATCTTTGGCGAGTCCGGGAGCCATAAGTGTTCCTAGCTTATATCTGTCTCCATGGC 797
QY 836 CAAAGGCTTATTCACAAAGCCATCATGAGTGGGGTGGCCATCATCCCTTACCTGGA 895
DB 798 CAAAGGCTTATTCACAAAGCCATCATGAGTGGGGTGGCCATCATCCCTTACCTGGA 857
QY 896 GGGCCATGATTTAGGAAGAGTGAAGACTTGAGGTGTTGCAATTTCTGTGGTAACA 955
DB 858 GGGCCATGATTTAGGAAGAGTGAAGACTTGAGGTGTTGCAATTTCTGTGGTAACA 917
```

```
QY 956 TGTGTGAGCTTGTGAGGCGCTTGTAGGTGCTGTGAGGACAAAACCTTCAAGAGTGTCT 1015
DB 918 TGTGTGAGCTTGTGAGGCGCTTGTAGGTGCTGTGAGGACAAAACCTTCAAGAGTGTCT 977
QY 1016 GACCCCTGAGCGAAGAAACAAAGCTCTTTCATCTGAGATGTGTGATGTGCTTTCTTCA 1075
DB 978 GACCCCTGAGCGAAGAAACAAAGCTCTTTCATCTGAGATGTGTGATGTGCTTTCTTCA 1037
QY 1076 TTAGCTCTAGATCTATGTTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGT 1135
DB 1038 TTAGCTCTAGATCTATGTTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGT 1097
QY 1136 CAATTAACCAAGTGTGCTTCTGCTGCTATGAAAGAGGCTCTGAGATCTCAAGTGG 1195
DB 1098 CAATTAACCAAGTGTGCTTCTGCTGCTATGAAAGAGGCTCTGAGATCTCAAGTGG 1157
QY 1196 CTCACAAAGTCCCTTGGCCCTTCATCTGATCAAAACATCTCTGACATCCCGGCTCACTA 1255
DB 1158 CTCACAAAGTCCCTTGGCCCTTCATCTGATCAAAACATCTCTGACATCCCGGCTCACTA 1217
QY 1256 TTTGACCTTGTGTGCTAATGAATCTTCATGACAGACACTCCCTGACTGAATCCGAGA 1315
DB 1218 TTTGACCTTGTGTGCTAATGAATCTTCATGACAGACACTCCCTGACTGAATCCGAGA 1277
QY 1316 CAGTCTTCTGAGACTTGTCTTGGAGATGTGTTGTGTGTGCTCCCTGACTGATCAAGCTCG 1375
DB 1278 CAGTCTTCTGAGACTTGTGTGGAGATGTGTTGTGTGTGCTCCCTGACTGATCAAGCTCG 1337
QY 1376 ATATCACAGAGATCTGTGTGACCTGTCTATGATGATTTGGGACCGGCTCTCACTG 1435
DB 1338 ATATCACAGAGATCTGTGTGACCTGTCTATGATGATTTGGGACCGGCTCTCACTG 1397
QY 1436 CTTTGAAGACAGAAAGCGGCTTTTGTCAAAAGCCGACAGCTGATGAAGTCCGCTTGT 1495
DB 1398 CTTTGAAGACAGAAAGCGGCTTTTGTCAAAAGCCGACAGCTGATGAAGTCCGCTTGT 1457
QY 1496 GTTGTGTGTGCTCTTCTCTGAAGGGGAGCAATGTGTAATGTGGAAGGAGCCAGGAGAGA 1555
DB 1458 GTTGTGTGTGCTCTTCTCTGAAGGGGAGCAATGTGTAATGTGGAAGGAGCCAGGAGAGA 1517
QY 1556 GAAGTACTGAGCCGGAAGATGATGAATATCTGGGCTAACCCTTTGCTGAAACCGGAAATCC 1615
DB 1518 GAAGTACTGAGCCGGAAGATGATGAATATCTGGGCTAACCCTTTGCTGAAACCGGAAATCC 1577
QY 1616 TAAATGGGAAGCACTGTCTCTGTGGGCAAGCTTAAATCTGAAGAGATCACTCCAGCT 1675
DB 1578 TAAATGGGAAGCACTGTCTCTGTGGGCAAGCTTAAATCTGAAGAGATCACTCCAGCT 1637
QY 1676 GAACTTGAACATGAGCCCTCGGACAGAGACTCAAGAAACCGGAGTGAATTTTGAACAG 1735
DB 1638 GAACTTGAACATGAGCCCTCGGACAGAGACTCAAGAAACCGGAGTGAATTTTGAACAG 1697
QY 1736 CACCATCCCTCTGATCTGTGTGCTGCTCGACATGCTCACAAGTCTCTTCTTTAAAC 1795
DB 1698 CACCATCCCTCTGATCTGTGTGCTGCTCGACATGCTCACAAGTCTCTTCTTTAAAC 1757
QY 1796 TTTCTCTCTCTCTCCAGGCTTCTTTCTTTTGTGAGCTCCCTGAGAGATTAATCTTCT 1855
DB 1758 TTTCTCTCTCTCTCCAGGCTTCTTTCTTTTGTGAGCTCCCTGAGAGATTAATCTTCT 1817
QY 1818 GTGATTTTGTGTTCCCTTCTCTCTCTATATATTTCTCCGGAATCATTTAGCTTTCTGA 1877
DB 1856 GTGATTTTGTGTTCCCTTCTCTCTCTCAATATTTCTCCGGAATCATTTAGCTTTCTGA 1915
QY 1916 GCTCAGCTGCTTCTATAGGGAATCTTGCAAAACAGGCTGTTTGGCTGATATTTATAG 1975
DB 1878 GCTCAGCTGCTTCTATAGGGAATCTTGCAAAACAGGCTGTTTGGCTGATATTTATAG 1937
QY 1976 ACTTAGGAATGATCTTTACAGATCTTTTCAACATCAAAAGTGAATTTTGTCTTGA 2035
DB 1938 ACTTAGGAATGATCTTTACAGATCTTTTCAACATCAAAAGTGAATTTTGTCTTGA 1997
QY 2036 GGCACAAAGATTTCTTCAATTAATTTGGAAGAGGCTGACTTATGATTTGCTAATAAT 2095
```

|||||
Db 1998 GGCACACAAATTTCCCAATTAATTTGGAAGGGCTGGCTTATGTTGATATATAT 2057
Qy 2096 GGTTCATCATATGAAATTAATTCGAATGT 2130
Db 2058 GGTTCATCATATGAAATTAATTCGAATGT 2092
RESULT 6
AB086169
ID AB086169 standard; DNA; 1728 BP.
XX
XX AB086169;
AC
XX
XX 10-SEP-2002 (first entry)
XX
XX Novel human gene. SEQ ID 40.
DE
XX Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
XX necrotic; neuroprotective; immunosuppressive; haemostatic;
XX antiinflammatory; cardiant; antilicer; virucide; antithyroid;
XX cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
XX wound healing disorders; atherosclerosis; Parkinson's disease;
XX Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
XX inflammation; neoplastic disease; nervous system disorder;
XX cardiovascular disorder; pancreaticitis; respiratory disorder;
XX hyperproliferation; systemic autoimmune disease; hyper-immunity;
XX developmental abnormality; gastrointestinal ulceration; neuropathy;
XX haematological disease; metabolic disease; sperm dysfunction;
XX thyroid disorder; hypothyroidism; brain damage; colitis;
XX cone photo-transduction deficiency; neurological disease; stroke;
XX anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
XX trachea; thymus; lymph node; muscular system; obesity; anorexia;
XX growth abnormality; precocious puberty; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200250105-A1.
PN
XX 27-JUN-2002.
PD
XX 17-DEC-2001; 2001WO-US049232.
PF
XX 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX (SMK) SMITHKLINE BEECHAM CORP.
XX (SMK) SMITHKLINE BEECHAM PLC.
XX (GLAX) GLAXO GROUP LTD.
XX
XX Aggarwal P, Birkeland M, Cogswell JP, Kahnlick KF, Lai Y;
XX Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX WPI; 2002-508784/54.
XX P-PSDB; ABP61004.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
XX or treating various diseases e.g. cancer, wounds, atherosclerosis,
XX Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX Claim 2(a); Page 251; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
XX which allow it to be secreted extracellularly or membrane associated. The
XX activity of polypeptides of the invention may be described as,
XX cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, necrotic,
XX neuroprotective, immunosuppressive, haemostatic, antiinflammatory,

CC cardiant, antilicer, virucide, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreaticitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, anglogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records AB086130-AB086184 represent novel human CDNA's
CC of the invention
XX
XX
SQ Sequence 1728 BP; 367 A; 483 C; 437 G; 441 T; 0 U; 0 Other;
Query Match 76.6%; Score 1652.8; DB 6; Length 1728;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1654; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 186 GGGCTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGACAA 245
Db 73 GGGCTTCTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGACAA 132
Qy 246 GTGACTGTCTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTGTGCT 305
Db 133 GTGACTGTCTGGAGAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTGTGCT 192
Qy 306 CCCCCGCTGGGATCCCTGGATTTAGAAACCCGACGCTGATCCGCTGGGATTAATTG 365
Db 193 CCCCCGCTGGGATCCCTGGATTTAGAAACCCGACGCTGATCCGCTGGGATTAATTG 252
Qy 366 CGAGAGGACCACTCTTACCTTAATTTGTGCTCCAGAACTCAGATGCTGCTTATAT 425
Db 253 CGAGAGGACCACTCTTACCTTAATTTGTGCTCCAGAACTCAGATGCTGCTTATAT 312
Qy 426 CAACATCTCAAGGTGATTAACCCGAAATTCGAGTGTCAAGAGTGTCTTACTG 485
Db 313 CAACATCTCAAGGTGATTAACCCGAAATTCGAGTGTCAAGAGTGTCTTACTG 372
Qy 486 AACATCTTATGCGCTGCCACGCGCATACAGGCTCAAGCTCCCGTGTGCTGCTTC 545
Db 373 AACATCTTATGCGCTGCCACGCGCATACAGGCTCAAGCTCCCGTGTGCTGCTTC 432
Qy 546 CCAGAGGTCCTTCAAGACTGGCTCAGGCTCATCTTGTATGGGTGCGCCTGCTGCG 605
Db 433 CCAGAGGTCCTTCAAGACTGGCTCAGGCTCATCTTGTATGGGTGCGCCTGCTGCG 492
Qy 606 TATGAGAGCTGCTGTGTGTGCTGTCAATACCGGCTTGAATATTTGGTTTCTTAC 665
Db 493 TATGAGAGCTGCTGTGTGTGCTGTCAATACCGGCTTGAATATTTGGTTTCTTAC 552
Qy 666 ACATGGATCAGATGCTTCGCGGAACTGGGCTTCAAGAGACAGTGGCTCTGTTC 725
Db 553 ACATGGATCAGATGCTTCGCGGAACTGGGCTTCAAGAGACAGTGGCTCTGTTC 612
Qy 726 TGGGTCCAGAGAAATCATGAGTTCCTTGGTGGGACCCAGCTCTGTGACATCTTTGG 785
Db 613 TGGGTCCAGAGAAATCATGAGTTCCTTGGTGGGACCCAGCTCTGTGACATCTTTGG 672
Qy 786 GAGTCCGCGGAGCATTAAGTGTTCATGCTTATACTGTCTCCATGGCCAAAGCTTA 845
Db 673 GAGTCCGCGGAGCATTAAGTGTTCATGCTTATACTGTCTCCATGGCCAAAGCTTA 732
Qy 846 TTCACAAAGCATATGAGAGTGGGTGGCCATATCCCTTACTTGAAGCCCATAT 905
Db 733 TTCACAAAGCATATGAGAGTGGGTGGCCATATCCCTTACTTGAAGCCCATAT 792
Qy 906 TATGAGAGAGTGAAGACTGAGGTGTTGACATTTCTGTGTGAACATGCTGCAAG 965

```
Db 793 TATGAGAAAGTGGAGACCTGCAAGTGGTTCACATTTCTGTGTAACATGCGCTCAGC 852
Qy 966 TCTGAGGCGCCCTGCTGAGGTGCTTGAGGACAAAACCTTCCAAAGAGCTGTGACCTTCAGC 1025
Db 853 TCTGAGGCGCCCTGCTGAGGTGCTTGAGGACAAAACCTTCCAAAGAGCTGTGACCTTCAGC 912
Qy 1026 CAGAAAACAAGTCTTTCACCTCGAGTGTGAGTGTGCTTCTTCTTAATGAGCCTCTA 1085
Db 913 CAGAAAACAAGTCTTTCACCTCGAGTGTGAGTGTGCTTCTTCTTAATGAGCCTCTA 972
Qy 1086 GATCTATTGTCTCAGAAAAGATTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1145
Db 973 GATCTATTGTCTCAGAAAAGATTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1032
Qy 1146 GAGTGTGAGTCTTCTGCTGCTTGAAGAGAGCTCTCTGAGTCTTCAAGTGTCTCAACAG 1205
Db 1033 GAGTGTGAGTCTTCTGCTGCTTGAAGAGAGCTCTCTGAGTCTTCAAGTGTCTCAACAG 1092
Qy 1206 TCCCTTGCCCTCCATCTGATACAAAACATCTGCAACATCCCGGCTGAGTATTTGCACTT 1265
Db 1093 TCCCTTGCCCTCCATCTGATACAAAACATCTGCAACATCCCGGCTGAGTATTTGCACTT 1152
Qy 1266 GTGGCTTAATGAATTAATCTTCCATGACAGCACTCCCTGACTGAATCCGAGACAGTCTTCTG 1325
Db 1153 GTGGCTTAATGAATTAATCTTCCATGACAGCACTCCCTGACTGAATCCGAGACAGTCTTCTG 1212
Qy 1326 GACTTGCTTGGAGATGTGTCTTGTGTGTGTCCTGCACTGATCACTGATGATTCACAA 1385
Db 1213 GACTTGCTTGGAGATGTGTCTTGTGTGTGTCCTGCACTGATGATGATTCACAA 1272
Qy 1386 GATGCTGTGTGCACTGTCTACTTCTATGAGTTCGCGACCGGCTGAGTCTTGAAGAC 1445
Db 1273 GATGCTGTGTGCACTGTCTACTTCTATGAGTTCGCGACCGGCTGAGTCTTGAAGAC 1332
Qy 1446 AGGAAGCGGCTTTTGTCAAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCGTGT 1505
Db 1333 AGGAAGCGGCTTTTGTCAAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCGTGT 1392
Qy 1506 GCCTTCCTGAAGGGGACATTTGTTATGTTGGAAGGACCAAGGAGAGAGAAAGTTACTG 1565
Db 1393 GCCTTCCTGAAGGGGACATTTGTTATGTTGGAAGGACCAAGGAGAGAGAAAGTTACTG 1452
Qy 1566 AACCAGAGATGAATGAATTAATGAGGCTTACCTTGTGCAACCGGGAATCTTAATGGAAAC 1625
Db 1453 AACCAGAGATGAATGAATTAATGAGGCTTACCTTGTGCAACCGGGAATCTTAATGGAAAC 1512
Qy 1626 GACCTGTCTCTGTGTGCGCACTTATATCTGACTGAGCACTCAAGCTGGAATTGAAC 1685
Db 1513 GACCTGTCTCTGTGTGCGCACTTATATCTGACTGAGCACTCAAGCTGGAATTGAAC 1572
Qy 1686 ATGAGCCTGCGACAGAGACTCAAAAGAACCGGCGGTGATTTTGGACCAAGCAATCCCC 1745
Db 1573 ATGAGCCTGCGACAGAGACTCAAAAGAACCGGCGGTGATTTTGGACCAAGCAATCCCC 1632
Qy 1746 CTGATCTGTCTGTGCTCCGACATGCTCAAGTCTCTTCTCTTAACCTTCTCTCT 1805
Db 1633 CTGATCTGTCTGTGCTCCGACATGCTCAAGTCTCTCTTCTTAACCTTCTCTCT 1692
Qy 1806 CTCTCCAGCCTTCTTCTTCTTCTTGTGTGCTCTTA 1841
Db 1693 CTCTCCAGCCTTCTTCTTCTTGTGTGCTCTTA 1728
```

RESULT 7
ADA53168
ID ADA53168 standard; cDNA; 1962 BP.
XX
AC ADA53168;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human coding sequence, SEQ ID 736.

```
XX KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
KW KM Gene therapy; human; secretory protein; membrane proteins; cancer;  
XX KM Inflammatory disease; osteoporosis; neurological disease; gene; ss.  
OS Homo sapiens.  
XX EP1293569-A2.  
PN 19-MAR-2003.  
PD 21-MAR-2002; 2002EP-00006586.  
XX PF 14-SEP-2001; 2001JP-00328381.  
XX PR 24-JAN-2002; 2002US-0350435P.  
XX PA (HELI-) RES ASSOC BIOTECHNOLOGY.  
XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX PI Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y,  
XX MPI: 2003-395539/38.  
DR P-PSDB; ADA54807.  
XX PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.  
XX PS Claim 1; SEQ ID NO 736; 205pp; English.  
XX CC The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.  
SQ Sequence 1962 BP; 428 A; 541 C; 473 G; 520 T; 0 U; 0 Other;  
Query Match 73.2%; Score 1579.4; DB 10; Length 1962;  
Best Local Similarity 92.0%; Pred. No. 0;  
Matches 1740; Conservative 0; Mismatches 1; Indels 150; Gaps 1;  
Qy 186 GGGCTTCTGTGAAGGCCACAGAGAACACCAAGCTGGAGTGAATTCAGGCAAGCAA 245  
Db 222 GGGCTTCTGTGAAGGCCACAGAGAACACCAAGCTGGAGTGAATTCAGGCAAGCAA 281  
Qy 246 GTCACTGTGTGGAAGCCCTGTGCTGTGAACGTGTTCCTCGGAAGTCCCTTGTGCTGT 305  
Db 282 GTCACTGTGTGGAAGCCCTGTGCTGTGACGTGTTCCTCGGAAGTCCCTTGTGCTGT 341  
Qy 306 CCCCCGCTGGAGTCCCTGCGATTTACGAACCCGACGCTGCATCGCCTGGATTAACCTTG 365  
Db 342 CCCCCGCTGGAGTCCCTGCGATTTACGAACCCGACGCTGCATCGCCTGGATTAACCTTG 401  
Qy 366 CGAGAAGCACTCTCAACCTTAATTTGTGTGCTCAAGAACTCAAGTGTGCTCTTAAGT 425  
Db 402 CGAGAAGCACTCTCAACCTTAATTTGTGTGCTCAAGAACTCAAGTGTGCTCTTAAGT 461  
Qy 426 CAACACATGTCTCAAGGTGATACCCGAATTCGAGGTGCAGAAAGTGCCTCTACCTG 485  
Db 462 CAACACATGTCTCAAGGTGATACCCGAATTCGAGGTGCAGAAAGTGCCTCTACCTG 521  
Qy 486 AACATCTATGCGCTGCGCCACGCGGATACAGGCTCCAAGTCCCGCTTGTGTGCTTC 545  
Db 522 AACATCTATGCGCTGCGCCACGCGGATACAGGCTCCAAGTCCCGCTTGTGTGCTTC 581  
Qy 546 CCAAGAGTGCCTTCAAGACTGCTCAGGCTCCATCTTGAATGGGTCGCGCTGCTGCTG 605  
Db 582 CCAAGAGTGCCTTCAAGACTGCTCAGGCTCCATCTTGAATGGGTCGCGCTGCTGCTG 641  
Qy 606 TATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
```


PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0289952P.
PR 19-JUN-2001; 2001US-0289237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.

(CURA-) CURAGEN CORP.

XX
XX
PI Guo X, Kekuda R, Miller CE, Malyanker UM, Spytek KA,
PI Paturajan M, Liu X, Gusev VY, La U, Vernet CM, Zerkhusen BD,
PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V,
PI Padigar M, Shinkels RA, Gangoli EA, Taupier RJ, Casman SJ, JI W,
PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ,
PI McDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA,
PI Ellerman K;
XX
XX

DR MPI; 2003-046858/04.
DR P-PSDB; ABUS4639.
XX
XX

PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.

XX
XX
PS Claim 17, Page 303; 666p; English.

XX
XX
CC The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, hemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABX72170-ABX72275 represent human NOVX
CC polynucleotides of the invention
XX
XX

Sequence 1746 BP; 402 A; 457 C; 459 G; 428 T; 0 U; 0 Other;

Query Match 66.5%; Score 1435.6; DB 8; Length 1746;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1996; Conservative 0; Mismatches 59; Indels 15; Gaps 1;

QY 176 GGGAAAAAAGTGGGCTTCTGTGAGAGGCAAGAGAAACACAGGCTGGGATGATTCA 235
Db 69 GGTTCAGAGAGGCTTCTGTGAGAGGCAAGAGAAACACAGGCTGGGATGATTCA 128
QY 236 GGGCAAGCAAGTCAATGTCGTCGGAGAGCCCTGTGCTGTGAACGTGTTCTCGAGTCC 295
Db 129 GGGCAAGCAAGTCAATGTCGTCGGAGAGCCCTGTGCTGTGAACGTGTTCTCGAGTCC 188
QY 296 CTTTGCTGCTCCCCGCTGGGATCCCTGGAATTACGAACCCGAGCTGGATGGCCTG 355
Db 189 CTTTGCTGCTCCCCGCTGGGATCCCTGGAATTACGAACCCGAGCTGGATGGCCTG 248
QY 356 GGAATTAATTTGCAAGAGCACTCTTACCTTAATTTGCTGCTCAAGAACTGAGTGGCT 415
Db 249 GGAATTAATTTGCAAGAGCACTCTTACCTTAATTTGCTGCTCAAGAACTGAGTGGCT 308

QY 416 GCTCTTAGATCAACATGCTCAAGTGATACCGGAAATTCGGAGTGTGAGAGACTG 475
Db 309 GCTCTTAGATCAACATGCTCAAGTGATACCGGAAATTCGGAGTGTGAGAGACTG 368
QY 476 CCTTACCTGAACATCTATAGCGCTTGGCCCAAGCGATACAGGCTCAAGCTCCCGCTT 535
Db 369 CCTTACCTGAACATCTATAGCGCTTGGCCCAAGCGATACAGGCTCAAGCTCCCGCTT 428
QY 536 GGTGTGTGTCACGAGAGGCTTCAAGATGAGCTCAAGCTTCAAGTGTGAGTCCG 595
Db 429 GGTGTGTGTCACGAGAGGCTTCAAGATGAGCTCAAGCTTCAAGTGTGAGTCCG 488
QY 596 CCGTGTGTGTCACGAGAGGCTTCAAGATGAGCTCAAGCTTCAAGTGTGAGTCCG 655
Db 489 CCGTGTGTGTCACGAGAGGCTTCAAGATGAGCTCAAGCTTCAAGTGTGAGTCCG 548
QY 656 TTTCTTCAACCAATGGATTCAGATGCTCCGAGGAACTGGGCTTCAAGAGCAAGTGGC 715
Db 549 TTTCTTCAACCAATGGATTCAGATGCTCCGAGGAACTGGGCTTCAAGAGCAAGTGGC 608
QY 716 TGCTCTGTCTGGGTCAGAGAAACATGAGTCTTGGTGAGGAGCCAGCTGTGAGC 775
Db 609 TGCTCTGTCTGGGTCAGAGAAACATGAGTCTTGGTGAGGAGCCAGCTGTGAGC 668
QY 776 CATCTTTGGCGAGTCCGCGGAGCCATTAAGTGTCTTAAGTCTTAAGTGTCTCCATGGC 835
Db 669 CATCTTTGGCGAGTCCGCGGAGCCATTAAGTGTCTTAAGTGTCTCCATGGC 728
QY 836 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTCATTCCTTACCTGCA 895
Db 729 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTCATTCCTTACCTGCA 788
QY 896 GAGCCATGATTTATGAGAAAGTGAAGACCTGAGAGTGTGACATTTCTGTGTAACA 955
Db 789 GAGCCATGATTTATGAGAAAGTGAAGACCTGAGAGTGTGACATTTCTGTGTAACA 848
QY 956 TGCGTCAGACTGTAGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1015
Db 849 TGCGTCAGACTGTAGGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 908
QY 1016 GACCTTGAAGCAAGAAACAAAGTCTTCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1075
Db 909 GACCTTGAAGCAAGAAACAAAGTCTTCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 968
QY 1076 TGACCTTGAATTTATGTCAGAAAGATTTAAAGCAATTCCTTCATCATGAGT 1135
Db 969 TGACCTTGAATTTATGTCAGAAAGATTTAAAGCAATTCCTTCATCATGAGT 1028
QY 1136 CAATTAACCAAGAGTGTGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1195
Db 1029 CAATTAACCAAGAGTGTGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1088
QY 1196 CTCACAAAGTCTTGTGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1240
Db 1089 CACTCCCTCAACCGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1148
QY 1241 CATCCGCTCAATTTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1300
Db 1149 CATCCGCTCAATTTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1208
QY 1301 GACTGAATTCGAGAGATCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1360
Db 1209 GACTGAATTCGAGAGATCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1268
QY 1361 ACTGATCAAGCTGATATCAAGAGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1420
Db 1269 ACTGATCAAGCTGATATCAAGAGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1328
QY 1421 GCACGGGCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1480
Db 1329 GCACGGGCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1388

QY 1481 TGAAGTCCTGTTGTTGCGTGGTCCCTTCTGAAGGGGAGACATTGTATGTTGCAAGG 1540
DB 1389 TGAAGTCCTGTTGTTGTTGCGTGGTCCCTTCTGAAGGGGAGACATTGTATGTTGCAAGG 1448
QY 1541 AGCCACGAGAGAGAGAGATTACTGAGCCCGAAGATGATGAATACTGGGCTTACCTTTGC 1600
DB 1449 AGCCACGAGAGAGAGAGATTACTGAGCCCGAAGATGATGAATACTGGGCTTACCTTTGC 1508
QY 1601 TCGAACCGGGAATCTTAATGAGAACGACCTGTCTGTGGCCAGCTTATATCTGACTGA 1660
DB 1509 TCGAACCGGGAATCTTAATGAGAACGACCTGTCTGTGGCCAGCTTATATCTGACTGA 1568
QY 1661 GCAGTACCTCCAGCTGGAATTGAACATGAGCCTCGACAGAGACTCAAGAAACCGCGGGT 1720
DB 1569 GCAGTACCTCCAGCTGGAATTGAACATGAGCCTCGACAGAGACTCAAGAAACCGCGGGT 1628
QY 1721 GGATTTTGG 1730
DB 1629 AGATGTGTGG 1638

RESULT 9
ABQ86170
ID ABQ86170 standard; DNA; 1746 BP.
XX
AC ABQ86170;
XX
DT 10-SEP-2002 (first entry)
XX
DE Novel human gene. SEQ ID 41.
XX
XX Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
XX neurotropic; neuroprotective; immunosuppressive; haemostatic;
XX antiinflammatory; cardiac; antiviral; virucide; antithyroid;
XX cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
XX wound healing disorders; atherosclerosis; Parkinson's disease;
XX Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
XX inflammation; neoplastic disease; nervous system disorder;
XX cardiovascular disorders; pancreatitis; respiratory disorder;
XX hyperproliferation; systemic autoimmune disease; hyper-immunity;
XX developmental abnormality; gastrointestinal ulceration; neuropathy;
XX haematological disease; metabolic disease; sperm dysfunction;
XX chondro disorder; hypothyroidism; brain damage; colitis;
XX cone photo-transduction deficiency; neurological disease; stroke;
XX anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
XX trachea; thymus; lymph node; muscular system; obesity; anorexia;
XX growth abnormality; precocious puberty; gene; ss.
OS Homo sapiens.
XX
XX MO200250105-A1.
PN
XX
XX 27-JUN-2002.
PD
PF 17-DEC-2001; 2001MO-US049232.
XX
XX 19-DEC-2000; 2000US-0256710P.
XX 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
XX
XX Agatwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q,
XX
XX WPI; 2002-508784/54.

DR P-PSDB; ABP61005.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis;
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
PS Claim 2(a); Page 251-252; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiac, anticancer, virucide, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, anglogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
XX of the invention
SQ Sequence 1746 BP; 401 A; 464 C; 459 G; 422 T; 0 U; 0 Other;
XX
XX Query Match 66.4%; Score 1432.4; DB 6; Length 1746;
XX Best Local Similarity 95.4%; Pred. No. 0;
XX Matches 1494; Conservative 0; Mismatches 51; Indels 21; Gaps 1;
XX
QY 186 GGGCTTTGCTGAAAGGGCCACAGAGAAACACAGGCTGGGATGATTCAGGGCAAGCA 245
DB 73 GGGCTTTGCTGAAAGGGCCACAGAGAAACACAGGCTGGGATGATTCAGGGCAAGCA 132
QY 246 GTCAGTGTGTGGAAGCCCTGTGCTGGAAGTGTCTCTCGAGTCCCTTTGCTGT 305
DB 133 GTCAGTGTGTGGAAGCCCTGTGCTGGAAGTGTCTCTCGAGTCCCTTTGCTGT 192
QY 306 CCCCGCTGGATCCCTGCGATTACGAACCCGACGCTGATCGCCCTGGATTA 365
DB 193 CCCCGCTGGATCCCTGCGATTACGAACCCGACGCTGATCGCCCTGGATTA 252
QY 366 CGAAGAGCCACCTCCATACCTATTTGCTGCTCAAGACTCAAGTGGGCTTTAGAT 425
DB 253 CGAAGAGCCACCTCCATACCTATTTGCTGCTCAAGACTCAAGTGGGCTTTAGAT 312
QY 426 CAACATGCTCAAGGTGATTCACCGAAATTCGAGTGCAGAAAGTCTCTACTG 485
DB 313 CAACATGCTCAAGGTGATTCACCGAAATTCGAGTGCAGAAAGTCTCTACTG 372
QY 486 AACATTTATGCGCTGCGCCACCGCGATACAGGCTCAAGCTCCCGTGTGTGAGTT 545
DB 373 AACATTTATGCGCTGCGCCACCGCGATACAGGCTCAAGCTCCCGTGTGTGAGTT 432
QY 546 CGAAGAGTGCCTTCAAGACTGCTCAAGCTCAATTTTGAATGGGCTCCGCTGGCTGCC 605
DB 433 CGAAGAGTGCCTTCAAGACTGCTCAAGCTCAATTTTGAATGGGCTCCGCTGGCTGCC 492
QY 606 TATGAGAGCGTGTGTGTGTGCTGCTCAAGTCCGCTAGAAATTTATTTGTTCTTAC 665
DB 493 TATGAGAGCGTGTGTGTGTGCTGCTCAAGTCCGCTAGAAATTTATTTGTTCTTAC 552
QY 666 ACATGGGATCAGATCTCCGGGAACTGGGCTTCAAGAGCAGGATGCTGCTGTCC 725
DB 553 ACATGGGATCAGATCTCCGGGAACTGGGCTTCAAGAGCAGGATGCTGCTGTCC 612

726 TGGGTCGAGAGAAATGAGTCTTCCGTGGGAGCCCAAGCTGTGACATCTTTGGC 785
 613 TGGGTCGAGAGAAATGAGTCTTCCGTGGGAGCCCAAGCTGTGACATCTTTGGC 672
 786 GAGTCCGCGGAGGACATTAAGTCTTCTAGTCTTACTAGTCTCCCATGGCCAAAGGCTTA 845
 673 GAGTCCGCGGAGGACATTAAGTCTTCTAGTCTTACTAGTCTCCCATGGCCAAAGGCTTA 732
 846 TTCCCAAAAGCCATCATGAGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGAGCCCATGAT 905
 733 TTCCCAAAAGCCATCATGAGAGAGTGGGGTGGCCATCATCCCTTAACCTGAGAGCCCATGAT 792
 906 TATGAGAAAGTGAAGACCTGACAGTGGTTCACATTTCTGTGGTAAACAATGCGTCAAC 965
 793 TATGAGAAAGTGAAGACCTGACAGTGGTTCACATTTCTGTGGTAAACAATGCGTCAAC 852
 966 TCTGAGGCGCTGAGAGTGGCTTGAAGCAAAACCTTCAAGAGCTGTGACCCCTGAGC 1025
 853 TCTGAGGCGCTGAGAGTGGCTTGAAGCAAAACCTTCAAGAGCTGTGACCCCTGAGC 912
 1026 CAGAAAACAAAGTCTTCACTCGAGTGGTGAAGTGGTCTTCTTCTTAATGAGCTCTA 1085
 913 CAGAAAACAAAGTCTTCACTCGAGTGGTGAAGTGGTCTTCTTCTTAATGAGCTCTA 972
 1086 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCAATGAGGATCAATTAACAC 1145
 973 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCAATGAGGATCAATTAACAC 1032
 1146 GAGTGGGCTTCCGTGCTGCTTGAAGAGAGCTTCTGAGATCCTCAATGAGGCTTCAACAG 1205
 1033 GAGTGGGCTTCCGTGCTGCTTGAAGAGAGCTTCTGAGATCCTCAATGAGGCTTCAACAG 1092
 1206 TCCCTT-----GCCCTCATCTGATCAAAACATCTGACATC 1244
 1093 AACCGTATGACAGCTTTGGCTTCAACAGCTGGGCAATTCACAGAAAGATCAGCAATC 1152
 1245 CCGCCTCAGATTTGCACTTGTGGCTTAATGAATTAATCTTCATGACAGACATCCCTGACT 1304
 1153 CCGCCTCAGATTTGCACTTGTGGCTTAATGAATTAATCTTCATGACAGACATCCCTGACT 1212
 1305 GAAATCCGAGACAGTCTTGTGACCTTGTGAGAGATGTTGTTGGTCCCTGACATC 1364
 1213 GAAATCCGAGACAGTCTTGTGACCTTGTGAGAGATGTTGTTGGTCCCTGACATC 1272
 1365 ATCAGAGCTGATATCAACAGAGATGCTGTGCACTGTCTAATCTTATGATGTTGGGAC 1424
 1273 ATCAGAGCTGATATCAACAGAGATGCTGTGCACTGTCTAATCTTATGATGTTGGGAC 1332
 1425 CCGCCTCAGATTTGTAAGACAGAGCCGCTTTGTCAAAAGCCGACCAAGCTGATGA 1484
 1333 CCGCCTCAGATTTGTAAGACAGAGCCGCTTTGTCAAAAGCCGACCAAGCTGATGA 1392
 1485 GTCGCGTTTGTGTCGATGATGCTTCTGAAAGGGGACATTTATGTTTCGAGAGAGCC 1544
 1393 GTCGCGTTTGTGTCGATGATGCTTCTGAAAGGGGACATTTATGTTTCGAGAGAGCC 1452
 1545 ACGAGAGAGAGAAATGTAATGAGCCGGAAGATGTAATTAATCTGAGAGAGAG 1604
 1453 ACGAGAGAGAGAAATGTAATGAGCCGGAAGATGTAATTAATCTGAGAGAGAG 1512
 1605 ACGGGAATCTTAATGAGAGACAGCTGTCTGTGGCCAGCTTAATTAATCTGAGAGAG 1664
 1513 ACGGGAATCTTAATGAGAGACAGCTGTCTGTGGCCAGCTTAATTAATCTGAGAGAG 1572
 1665 TACCTTCAGCTGAGACTTGAACATGAGCTTCGACAGAGACTTCAAGAGACCGCGGATGAT 1724
 1573 TACCTTCAGCTGAGACTTGAACATGAGCTTCGACAGAGACTTCAAGAGACCGCGGATGAT 1632
 1725 TTTTGG 1730
 1633 GTGTGG 1638

RESULT 10
 ADR19681
 ID ADR19681 standard; DNA; 1857 BP.
 XX
 AC ADR19681;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Human drug metabolizing enzyme (DME)-2 gene sequence.
 XX
 KW drug metabolizing enzyme; DME; cytostatic; immunosuppressive;
 KW anti-inflammatory; endocrine; ophthalmological; gastrointestinal;
 KW hepatocytic; cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; endocrine disorder; eye disorder;
 KW gastrointestinal disorder; liver disorder; metabolic disorder; gene; ds;
 human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1857
 FT /tag=a
 FT /product="Human drug metabolizing enzyme (DME) 2"
 XX
 PN MO200226988-A2.
 XX
 PD 04-APR-2002.
 XX
 PE 28-SEP-2001; 2001WO-US030662.
 XX
 PR 29-SEP-2000; 2000US-0236947P.
 PR 06-OCT-2000; 2000US-0238864P.
 PR 20-OCT-2000; 2000US-0243232P.
 PR 09-NOV-2000; 2000US-0247581P.
 PR 16-NOV-2000; 2000US-0249519P.
 PR 22-NOV-2000; 2000US-0252834P.
 PR 30-NOV-2000; 2000US-0250567P.
 XX
 PA (INCY)- INCYTE GENOMICS INC.
 XX
 PI Azimzal Y, Baughin MR, Borowsky ML, Ding L, Dugan BM,
 PI Elliott VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan PA;
 PI Lal P, Lee BA, Lu DM, Nguyen DB, Arvizu C, Pollicky JJ, Rankumar J;
 PI Ring RZ, Sanjanwala MS, Tang YF, Tribouley CM, Narinder WK;
 PI Walsh RT, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
 XX
 WP1: 2002-362498/39.
 DR P-PSDB; ADR19683.
 XX
 PT Human drug metabolizing enzymes, useful in the diagnosis and treatment of
 PT disorders associated with aberrant (DME) activity, e.g., cancer and
 PT autoimmune disorders.
 XX
 PS Claim 12; SEQ ID NO 20; 142bp; English.
 XX
 CC This invention relates to novel drug metabolizing enzymes (DME) and the
 CC nucleotide sequences which encode them. The invention may be useful for
 CC the development of compounds with a cytostatic, immunosuppressive,
 CC anti-inflammatory, endocrine, ophthalmological, gastrointestinal or
 CC hepatocytic activity acting as an agonist or antagonist of drug
 CC metabolizing enzyme activity. The invention may be used in the diagnosis
 CC and treatment of disorders associated with decreased or increased
 CC expression or activity of drug metabolizing enzymes. Such disorders
 CC include cancer, cell proliferative disorders, autoimmune/inflammatory,
 CC endocrine, eye, gastrointestinal (including liver disorders) and
 CC metabolic disorders. The present sequence is that of a gene which encodes
 CC a human drug metabolizing enzyme (DME) of the invention. Note: This
 CC sequence did not form part of the printed specification but was obtained
 CC in electronic format from EPO.
 XX
 SQ Sequence 1857 BP; 428 A; 487 C; 496 G; 446 T; 0 U; 0 Other;

Query Match 64.2%; Score 1386.2; DB 7; Length 1857;

Best Local Similarity 94.7%; Pred. No. 0;
Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

```

QY 176 GGGAAAACTGGGCGCTTCTGCTGAAGGCGCACAGAGAAACACAGCGCTGGAGATGATTC 235
Db 264 GGTCCAGAGAGGGCTTCTGTCTGAAGGGCCACAGAGAAACACAGCGTGGAGATGATTC 323
QY 236 GGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTCTCGAGATCCC 295
Db 324 GGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTGCTGTGAAGTGTCTCTCGAGATCCC 383
QY 296 CTCTTGTGCTCCCGCTGGGATCCCTGCGATTTACGAACCCGACGCTGCACTGCGCTG 355
Db 384 CTCTTGTGCTCCCGCTGGGATCCCTGCGATTTACGAACCCGACGCTGCACTGCGCTG 443
QY 356 GGAATACTTGGGAAAGCCCTCTACCTATTTGTGCTTCAAGATCTCAAGATGAGCT 415
Db 444 GGAATACTTGGGAAAGCCCTCTACCTATTTGTGCTTCAAGATCTCAAGATGAGCT 503
QY 416 GCTCTTAGATCAACATGCTCAAGGTGATACCCGAATTCGAGTGTCAAGAGACTG 475
Db 504 GCTCTTAGATCAACATGCTCAAGGTGATACCCGAATTCGAGTGTCAAGAGACTG 563
QY 476 CCTTACCTGAACATCTATGCGCTGCGCCACCGCGATACAGGCTCCAGCTCCCGCTT 535
Db 564 CCTTACCTGAACATCTATGCGCTGCGCCACCGCGATACAGGCTCCAGCTCCCGCTT 623
QY 536 GGTGTGTGTTCCAGAGAGTGCCTTCAAGATGAGCTCAAGCTCCATCTTTGATGGTCCG 595
Db 624 GGTGTGTGTTCCAGAGAGTGCCTTCAAGATGAGCTCAAGCTCCATCTTTGATGGTCCG 683
QY 596 CCGGCTGCTTGAAGAGCGTGTGTTGTGTGTGCTCAAGTACCGGCTGGAATATTTGG 655
Db 684 CCGGCTGCTTGAAGAGCGTGTGTTGTGTGTGCTCAAGTACCGGCTGGAATATTTGG 743
QY 656 TTTCTTCAACACATGGGATCAAGCATCTCCGAGAACTGGGCTTCAAGACCAAGTGGC 715
Db 744 TTTCTTCAACACATGGGATCAAGCATCTCCGAGAACTGGGCTTCAAGACCAAGTGGC 803
QY 716 TGTCTGTCTGTGGGTCAAGAAACATCGAGTTCTTGGTGGGAGCCCAAGCTGTGTAC 775
Db 804 TGTCTGTCTGTGGGTCAAGAAACATCGAGTTCTTGGTGGGAGCCCAAGCTGTGTAC 863
QY 776 CATCTTGGCGAGTCGCGGAGGCAATAAGTTTCTAGTCTTATCTGTCTCCATGGC 835
Db 864 CATCTTGGCGAGTCGCGGAGGCAATAAGTTTCTAGTCTTATCTGTCTCCATGGC 923
QY 836 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCTTACCTGGA 895
Db 924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCTTACCTGGA 983
QY 896 GGCCCATGATATGAGAAAGTGAAGACCTGACAGGTGTTGCATTTCTGTGTAAACA 955
Db 984 GGCCCATGATATGAGAAAGTGAAGACCTGACAGGTGTTGCATTTCTGTGTAAACA 1043
QY 956 TGTGCTCAAGCTGAGGCGCTGTGAGGTGCTGAGAGAAACCTTCCAAAGAGCTGCT 1015
Db 1044 TGTGCTCAAGCTGAGGCGCTGTGAGGTGCTGAGAGAAACCTTCCAAAGAGCTGCT 1103
QY 1016 GACCCCTCAGCCAGAAACAAAGCTTTCATCGAGTGTGTGATGCTTTCTTCTTAA 1075
Db 1104 GACCCCTCAGCCAGAAACAAAGCTTTCATCGAGTGTGTGATGCTTTCTTCTTAA 1163
QY 1076 TGAAGCTTAGATCTATTTGTCTCAGAAAGCATTTAAGCAATTCCTTCAATCGGAGT 1135
Db 1164 TGAAGCTTAGATCTATTTGTCTCAGAAAGCATTTAAGCAATTCCTTCAATCGGAGT 1223
QY 1136 CAATTAACAAGAGTGTGCTTCTGTGCTTGAAGAGAGGCTCCGAATCTCAAGTGG 1195
Db 1224 CAATTAACAAGAGTGTGCTTCTGTGCTTGAAGAGAGGCTCCGAATCTCAAGTGG 1255
QY 1196 CTCCAACAAGATCCCTTGCCTTCATCTGATACAAAACATCTGACATCCGCGCTCAGTA 1255

```

```

Db 1256 -----TGCATCCCGCTCAGTA 1274
QY 1256 TTTGCACTTGTGGCTAATGAATACCTTCATGACAGACCTCCCTGACGTGAATCCGAG 1315
Db 1275 TTTGCACTTGTGGCTAATGAATACCTTCATGACAGACCTCCCTGACGTGAATCCGAG 1334
QY 1316 CAGTCTTCTGGAATTGCTTGGAGATGTGTTGTGTGCTTGTGCTGCACTGATCAAGCTCG 1375
Db 1335 CAGTCTTCTGGAATTGCTTGGAGATGTGTTGTGTGCTTGTGCTGCACTGATCAAGCTCG 1394
QY 1376 ATATCAACAAGATGTGTGTGACCTGTCTACTTATGAGATTTCGACACCGGCTCAATG 1435
Db 1395 ATATCAACAAGATGTGTGTGACCTGTCTACTTATGAGATTTCGACACCGGCTCAATG 1454
QY 1436 CTTTGAAGACAGAAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTTGT 1495
Db 1455 CTTTGAAGACAGAAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTTGT 1514
QY 1496 GTTCGGTGTGCTTCTCTGAAGGGGACATTTGTATGTTGGAAGAGCACGAGAGAGA 1555
Db 1515 GTTCGGTGTGCTTCTCTGAAGGGGACATTTGTATGTTGGAAGAGCACGAGAGAGA 1574
QY 1556 GAAATTACTGAGCCGGAAGATGATGAATATCTGGGCTACCTTTGCTGAAACGGGAATCC 1615
Db 1575 GAAATTACTGAGCCGGAAGATGATGAATATCTGGGCTACCTTTGCTGAAACGGGAATCC 1634
QY 1616 TAATGGGAACGACCTGTCTGTGGCAGCTTATATGACTGACAGCACTACCTCAGCT 1675
Db 1635 TAATGGGAACGACCTGTCTGTGGCAGCTTATATGACTGACAGCACTACCTCAGCT 1694
QY 1676 GGACTTGAACATGAGCGCTCGACAGAGACTCAAGAACCAGCGGAGGATTTTGG 1730
Db 1695 GGACTTGAACATGAGCGCTCGACAGAGACTCAAGAACCAGCGGAGAGTGTGG 1749

RESULT 11
ADCS5523
ID ADCS5523 standard; cDNA; 1244 BP.
XX
AC ADCS5523;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human carboxylatase 24.64 encoding sequence.
XX
KW human carboxylatase-24.64; primary hypertension; digestive ulcer;
KW nephrotic; bronchial asthma; gene; ss.
XX
OS Homo sapiens.
XX
FT Key Location/Qualifiers
FT CDS 236..910
FT /*tag= a
XX
PN CN1382799-A.
XX
PD 04-DEC-2002.
XX
PF 26-APR-2001; 2001CN-00112736.
XX
PR 26-APR-2001; 2001CN-00112736.
XX
PA (BIOM-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI: 2003-269506/27.
XX
DR P-PsDB; ADCS5524.
XX
PS Polypeptide-human carboxylatase-24.64 and polynucleotide for coding it.
XX
Claim 6; SEQ ID NO 1; 31bp; Chinese.
XX

```

CC The present invention relates to a polypeptide-human carboxylase-24.64,
CC the polynucleotide for coding it, the process for preparing the
CC polypeptide by DNA recombination, the application of the polypeptide in
CC treating diseases such as primary hypertension, digestive ulcer,
CC nephrotic, bronchial asthma, tremor, etc, the antagonist of the polypeptide
CC and its medical action, and the application of the polynucleotide are
CC new. The present sequence represents human carboxylase 24.64 encoding
CC sequence.
XX
SQ Sequence 1244 BP, 301 A, 314 C, 271 G, 358 T, 0 U, 0 Other;

Query Match 56.3%; Score 1215; DB 10; Length 1244;
Best Local Similarity 99.6%; Pred. No. 1.2e-300;
Matches 1218; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 932 GGTTCGACATTTCTGTGTGTAACAATGCTGACACTGAGGCCCTGTGAGTGCCTGAG 991
DB 1 GGTTCGACATTTCTGTGTGTAACAATGCTGACACTGAGGCCCTGTGAGTGCCTGAG 60
QY 992 GACAAAACCTCCAGAGAGCTGACCTCAGCCAGAGAAAACAAGCTTTCACCTGAGT 1051
DB 61 GACAAAACCTCCAGAGAGCTGACCTCAGCCAGAGAAAACAAGCTTTCACCTGAGT 120
QY 1052 GGTTCGAGTGTCTTTCTTTCTTAATGAGCCTTAGATCTATGTCAGAAACATTTAA 1111
DB 121 GGTTCGAGTGTCTTTCTTTCTTAATGAGCCTTAGATCTATGTCAGAAACATTTAA 180
QY 1112 AGCAATTCCTTCATCATGAGAGTCAATTAACAAGTGTGGTCTCTGCTGCTTGA 1171
DB 181 AGCAATTCCTTCATCATGAGAGTCAATTAACAAGTGTGGTCTCTGCTGCTTGA 240
QY 1172 GAGAGCTCTGAGATCCTCAGTGGCTCAACAAGTCCCTGACCTCATCTGATACAAA 1231
DB 241 GAGAGCTCTGAGATCCTCAGTGGCTCAACAAGTCCCTGACCTCATCTGATACAAA 300
QY 1232 CATCTCGACATCCCGCTCAGTATTTGCACTTGTGGCTAATGAAATCTTCATGACA 1291
DB 301 CATCTCGACATCCCGCTCAGTATTTGCACTTGTGGCTAATGAAATCTTCATGACA 360
QY 1292 GCACTCCCTGACATGAAATCCAGAGAGCTCTTGTGAGCTTGGAGTGTCTTTGT 1351
DB 361 GCACTCCCTGACATGAAATCCAGAGAGCTCTTGTGAGCTTGGAGTGTCTTTGT 420
QY 1352 GGTCCCTGACATGATCAAGCTGATATCAAGAGTGTGTGACCTGTCTACTTCTA 1411
DB 421 GGTCCCTGACATGATCAAGCTGATATCAAGAGTGTGTGACCTGTCTACTTCTA 480
QY 1412 TGAATTTGGGACCCGGCTCAGTGTCTTGAAGACGAAAGCCGGCTTTTGTCAAGCCGA 1471
DB 481 TGAATTTGGGACCCGGCTCAGTGTCTTGAAGACGAAAGCCGGCTTTTGTCAAGCCGA 540
QY 1472 CCAGCTGATGAAGTCCGCTTTGTGTTCGATGTGCTTCTGAAAGGGGACATTTAT 1531
DB 541 CCAGCTGATGAAGTCCGCTTTGTGTTCGATGTGCTTCTGAAAGGGGACATTTAT 600
QY 1532 GTTCAAGAGAGCCAGAGAGAGAGAAATTAATGACCGGAAAGATGATGAATACTGAGC 1591
DB 601 GTTCAAGAGAGCCAGAGAGAGAGAAATTAATGACCGGAAAGATGATGAATACTGAGC 660
QY 1592 TACCTTTGCTCGAACCGGGAATCTTAATGGAAACGACCTGTCTGTGGCCAGCTTA 1651
DB 661 TACCTTTGCTCGAACCGGGAATCTTAATGGAAACGACCTGTCTGTGGCCAGCTTA 720
QY 1652 TCTGACTGAGAGTACTTCAGCTGAACTTGAACATGAGCCTCGAGACAGACTCAAGA 1711
DB 721 TCTGACTGAGAGTACTTCAGCTGAACTTGAACATGAGCCTCGAGACAGACTCAAGA 780
QY 1712 ACCGGGGGAGATTTTGGACACAGACATCCCTGATTCCTGTGCTCCGACATGCT 1771
DB 781 ACCGGGGGAGATTTTGGACACAGACATCCCTGATTCCTGTGCTCCGACATGCT 840
QY 1772 CCACAGTCTCTTTCTTCTTAATCTTCTCTCTCTCCAGCCTTTCTTTCTTTG 1831

DB 841 CCACAGTCTCTTTCTTCTTAATCTTCTCTCTCTCCAGCCTTTCTTTCTTTG 900
QY 1832 TGTCTCTTGAAGAGTATCTTCTGATTTGTTTCCCTCTCCCATTAATTTCTC 1891
DB 901 TGTCTCTTGAAGAGTATCTTCTGATTTGTTTCCCTCTCCCATTAATTTCTC 960
QY 1892 CCGCAATCTTAAGCTTCTTCTGAGCTCAGCTCTTCTATGAGGATCTTCAACAA 1951
DB 961 CCGCAATCTTAAGCTTCTTCTGAGCTCAGCTCTTCTATGAGGATCTTCAACAA 1020
QY 1952 GGTGCTTGTGCTGATTTTAATGACTTAAGAAATGATCTTAAGAAATTTCTTCAACAT 2011
DB 1021 GGTGCTTGTGCTGATTTTAATGACTTAAGAAATGATCTTAAGAAATTTCTTCAACAT 1080
QY 2012 CAAAAGTGCATTTGTCTGGAAGCAACAAGATTTTCTCAATTAATTTGGAAGGGC 2071
DB 1081 CAAAAGTGCATTTGTCTGGAAGCAACAAGATTTTCTCAATTAATTTGGAAGGGC 1140
QY 2072 TGGCTTAATTAAGTGTCAATTAATGATTTTGTAACTCATATGAATAATCAGATGTA 2131
DB 1141 TGGCTTAATTAAGTGTCAATTAATGATTTTGTAACTCATATGAATAATCAGATGTA 1200
QY 2132 AATAGAAAAA 2154
DB 1201 AATAGAAAAA 1223

RESULT 12
ADFS0146
ID ADFS0146 standard; cDNA; 2145 BP.
XX
AC ADFS0146;
XX
DT 12-FEB-2004 (first entry)
XX
DE Cat cauxin encoding cDNA SEQ ID NO:3.
XX
KW cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.
XX
OS Felis catus.
XX
PN JP2003250575-A.
XX
PD 09-SEP-2003.
XX
PE 04-MAR-2002; 2002JP-00057908.
XX
PR 04-MAR-2002; 2002JP-00057908.
XX
PA (TOHO-) TOHOKU TECHNOARCH KK.
XX
DR WPI; 2004-002277/01.
XX
DR P-PSDB; ADFS0147.
XX
PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX and for diagnosing cat kidney disease.
XX
PS Claim 6; SEQ ID NO 3; 33pp; Japanese.
XX
CC The present sequence encodes a cat cauxin protein (I) or its salt, which
CC is cat kidney disease marker. Also described: (1) a partial peptide (II)
CC of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
CC comprising (III); (4) a transformed host (V) comprising (III) or (IV);
CC (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
CC couples specifically with (I) or (II); (7) diagnosing cat kidney disease
CC which involves measuring (I) quantitatively, and where reduction of
CC amount of (I) indicates presence of the disease; (8) a cat kidney disease
CC diagnostic agent comprising (I) labelling agent; a reagent which measures
CC the biological activity of urinary (I) or (VI); and (9) a cauxin
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces

CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.

XX Sequence 2145 BP; 444 A; 581 C; 550 G; 564 T; 0 U; 6 Other;

Query Match 53.6%; Score 1156; DB 12; Length 2145;
Best Local Similarity 78.2%; Pred. No. 2.1e-285;
Matches 1466; Conservative 0; Mismatches 386; Indels 22; Gaps 6;

```
QY 186 GGGCCCTTGTGAGAGGGCCACAGAGAAACACAGCTGGATGTGATTGAGGCAAGCA 245
DB |||||
QY 228 GGGCCAGCTGTGATGACACAGAGAGACACAGGAGGATGGATCCGGGGAAGCA 297
DB GGGCCAGCTGTGATGACACAGAGAGACACAGGAGGATGGATCCGGGGAAGCA 297
QY 246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAAGCTGTCTCGAGATCCCTTTGCTCT 305
DB GTCACTGTGCTGGAGAGCCCTGTGCTGTGAAGCTGTCTCGAGATCCCTTTGCTCT 305
QY 298 ACCACTGTACTGGAGAGACCGTGTGTGAACATTTCTCGGATCCCTTATGTCTGA 357
DB ACCACTGTACTGGAGAGACCGTGTGTGAACATTTCTCGGATCCCTTATGTCTGA 357
QY 306 CCCCCGCTGGATCCCTGTGATTTACGAACCCGACAGCTGTGATCGCCCTGGATTA 365
DB CCCCCGCTGGATCCCTGTGATTTACGAACCCGACAGCTGTGATCGCCCTGGATTA 365
QY 358 CCTCTCTAGGGCCCTGTGATTTAGCAACAAAGCTGTCTGTCCGGGAATGACTTC 417
DB CCTCTCTAGGGCCCTGTGATTTAGCAACAAAGCTGTCTGTCCGGGAATGACTTC 417
QY 366 CGAGAGCCACTCTCTAACCCTTAATTTGTGCTTCAAGACTCAAGTGTGCTTAT 425
DB CGAGAGCCACTCTCTAACCCTTAATTTGTGCTTCAAGACTCAAGTGTGCTTAT 425
QY 418 CGAAATGCCAATCCCTAACCTTAATTTATCTTCAGAGACTTGAAGTGTGCTCT 477
DB CGAAATGCCAATCCCTAACCTTAATTTATCTTCAGAGACTTGAAGTGTGCTCT 477
QY 426 CAACAATGCTGAAGGCTAATTAACCCGAATTCGAGTGTCAAGAGACTGCTCTA 485
DB CAACAATGCTGAAGGCTAATTAACCCGAATTCGAGTGTCAAGAGACTGCTCTA 485
QY 478 CAACAGCTTCTCAAAATGTGCTTAACCCCAATTTGAAAGCTTCGAAACATGCT 537
DB CAACAGCTTCTCAAAATGTGCTTAACCCCAATTTGAAAGCTTCGAAACATGCT 537
QY 486 AACATCTATGCGCTGTCCACGACGATAGAGGCTCCAGCTCCCTCTTGTGTGTTC 545
DB AACATCTATGCGCTGTCCACGACGATAGAGGCTCCAGCTCCCTCTTGTGTGTTC 545
QY 538 AACATCTATGCGCTGTCCACGACGATAGAGGCTCCAGCTCCCTCTTGTGTGTTC 597
DB AACATCTATGCGCTGTCCACGACGATAGAGGCTCCAGCTCCCTCTTGTGTGTTC 597
QY 546 CCAGAGAGTGTCTTCAAGCTGTGCTCAAGCTCTCATTTTGAATGGTGTGCTGTG 605
DB CCAGAGAGTGTCTTCAAGCTGTGCTCAAGCTCTCATTTTGAATGGTGTGCTGTG 605
QY 598 CCGGGGGGCTCTTCAAGATGGGCTCAAGCTCTCTCTTGTGATGGTGTGCTGTG 657
DB CCGGGGGGCTCTTCAAGATGGGCTCAAGCTCTCTCTTGTGATGGTGTGCTGTG 657
QY 606 TATGAGAGCTGTGCTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 665
DB TATGAGAGCTGTGCTGTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 665
QY 658 TACGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 717
DB TACGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 717
QY 666 ACATGAGATCAGATGTCTCCGGGGAACCTGGGCTTTCAGAGCAAGTGTGCTGTG 725
DB ACATGAGATCAGATGTCTCCGGGGAACCTGGGCTTTCAGAGCAAGTGTGCTGTG 725
QY 718 ACAGGGGATGACATGTCCCGGGGGAACCTGGGCTTTCAGAGCAAGTGTGCTGTG 777
DB ACAGGGGATGACATGTCCCGGGGGAACCTGGGCTTTCAGAGCAAGTGTGCTGTG 777
QY 726 TGGGTTCAGAGAAATCAGATTTCTTGGTGGGGAACCCAGCTGTGTGACATTTTGG 785
DB TGGGTTCAGAGAAATCAGATTTCTTGGTGGGGAACCCAGCTGTGTGACATTTTGG 785
QY 778 TGGGTTCAGAGAAATCAGATTTCTTGGTGGGGAACCCAGCTGTGTGACATTTTGG 837
DB TGGGTTCAGAGAAATCAGATTTCTTGGTGGGGAACCCAGCTGTGTGACATTTTGG 837
QY 786 GAGTCCGGGGGACATTAAGTGTCTTACTTATATCTGTCTTCCCAATGGCCAAAGCT 845
DB GAGTCCGGGGGACATTAAGTGTCTTACTTATATCTGTCTTCCCAATGGCCAAAGCT 845
QY 838 GAGTCCGGGGGACATTAAGTGTCTTACTTATATCTGTCTTCCCAATGGCCAAAGCT 897
DB GAGTCCGGGGGACATTAAGTGTCTTACTTATATCTGTCTTCCCAATGGCCAAAGCT 897
QY 846 TTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACCTGTGAG--GCCAT 902
DB TTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACCTGTGAG--GCCAT 902
QY 898 TTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACCTGTGAG--GCCAT 957
DB TTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACCTGTGAG--GCCAT 957
QY 903 GATTATGAGAGAGTGAAGCTGTGAGTGTGTGCAATTTCTGTGTGAAATGTGCTGA 962
DB GATTATGAGAGAGTGAAGCTGTGAGTGTGTGCAATTTCTGTGTGAAATGTGCTGA 962
QY 958 GGTGATGAGAGAGAGATTTTGCAGTGTGTGACCGCTATCTGTGTGTCATGTGCT 1017
DB GGTGATGAGAGAGAGATTTTGCAGTGTGTGACCGCTATCTGTGTGTCATGTGCT 1017
QY 963 GACTCTGAGGCGCTGTGAGTGTGCTGTGAGAGCAAAACCTTCAAGAGTGTGACCTTC 1022
DB GACTCTGAGGCGCTGTGAGTGTGCTGTGAGAGCAAAACCTTCAAGAGTGTGACCTTC 1022
QY 1018 GACTCTGAGGCGCTGTGAGTGTGCTGTGAGAGCAAAACCTTCAAGAGTGTGACCTTC 1077
DB GACTCTGAGGCGCTGTGAGTGTGCTGTGAGAGCAAAACCTTCAAGAGTGTGACCTTC 1077
QY 1023 AGCCAGAAAACAAAGCTTCTCACTCGAGTGTGTGATGTGTCTTCTTCTAATGAGCT 1082
DB AGCCAGAAAACAAAGCTTCTCACTCGAGTGTGTGATGTGTCTTCTTCTAATGAGCT 1082
QY 1078 AGCAAGAAATCAAGTTCATTCAGATTTGATGATTTTCTTCTGTGATGAGCT 1137
DB AGCAAGAAATCAAGTTCATTCAGATTTGATGATTTTCTTCTGTGATGAGCT 1137
QY 1083 CTAGATCTAATGTCTCAGAAAGATTTAAAGCAATTCCTTCAATCAATGAGCTAAT 1142
DB CTAGATCTAATGTCTCAGAAAGATTTAAAGCAATTCCTTCAATCAATGAGCTAAT 1142
QY 1138 GTAGCCCTAATGACTCAAAAGCAATTTAATGATTCCTTCTATCATCGAGTCAATTA 1197
DB GTAGCCCTAATGACTCAAAAGCAATTTAATGATTCCTTCTATCATCGAGTCAATTA 1197
```

```
QY 1143 CACGAGTGTGCTTCTGTGCTGTGCTGTGAGAGAGCTCTGAGATCTCTAGTGTCCAC 1202
DB |||||
QY 1198 CACGAGTGTGCTTCTGTGCTGTGCTGTGAGAGAGCTCTGAGATCTCTAGTGTCCAC 1254
DB CACGAGTGTGCTTCTGTGCTGTGCTGTGAGAGAGCTCTGAGATCTCTAGTGTCCAC 1254
QY 1203 AAGTCCCTTGGCTCATGTGATCAAAACATCTGACATCTCCGCTCAATTTTGGAC 1262
DB AAGTCCCTTGGCTCATGTGATCAAAACATCTGACATCTCCGCTCAATTTTGGAC 1262
QY 1255 AAGTCTGTGCTGTGCTGTGATGACACGCTTGTGAATTTTCCACCAATTTTGGAC 1314
DB AAGTCTGTGCTGTGCTGTGATGACACGCTTGTGAATTTTCCACCAATTTTGGAC 1314
QY 1263 CTTTGGCTTAATGAAATTTCTCATGACAGCACTCCCTTACTGAAATTCGAGACATCT 1322
DB CTTTGGCTTAATGAAATTTCTCATGACAGCACTCCCTTACTGAAATTCGAGACATCT 1322
QY 1315 CTTTGGCTTAATGAAATTTCTCATGACAGCACTCCCTTACTGAAATTTTGGAC 1374
DB CTTTGGCTTAATGAAATTTCTCATGACAGCACTCCCTTACTGAAATTTTGGAC 1374
QY 1323 CTGAGCTTCTGTGAGATGTGTCTTGTGTGCTTCCCTGACACTGATCAAGCTGATTCAC 1382
DB CTGAGCTTCTGTGAGATGTGTCTTGTGTGCTTCCCTGACACTGATCAAGCTGATTCAC 1382
QY 1375 CTGAGCTTCTGTGAGATGTGTCTTGTGTGCTTCCCTGACACTGATTCAT 1434
DB CTGAGCTTCTGTGAGATGTGTCTTGTGTGCTTCCCTGACACTGATTCAT 1434
QY 1383 AAGATGTGTGTGACCTGTCTACTTATGAGTGTGAGACCGGCTCAAGTCTTGA 1442
DB AAGATGTGTGTGACCTGTCTACTTATGAGTGTGAGACCGGCTCAAGTCTTGA 1442
QY 1435 AAGATGTGTGTGACCTGTCTACTTATGAGTGTGAGACCGGCTCAAGTCTTGA 1494
DB AAGATGTGTGTGACCTGTCTACTTATGAGTGTGAGACCGGCTCAAGTCTTGA 1494
QY 1443 GACAGAGAGCGGCTTGTGCAAGCCGACAGGCTGATGAAGTGTGCTTGTGCT 1502
DB GACAGAGAGCGGCTTGTGCAAGCCGACAGGCTGATGAAGTGTGCTTGTGCT 1502
QY 1495 GACAGAGAGCGGCTTGTGCAAGCCGACAGGCTGATGAAGTGTGCTTGTGCT 1554
DB GACAGAGAGCGGCTTGTGCAAGCCGACAGGCTGATGAAGTGTGCTTGTGCT 1554
QY 1503 GGTGCTTCTGTGAGAGGAGACATTTGTATGTTGAGAGGACACGAGAGAGAGATTA 1562
DB GGTGCTTCTGTGAGAGGAGACATTTGTATGTTGAGAGGACACGAGAGAGAGATTA 1562
QY 1555 GGTGCTTCTGTGAGAGGAGACATTTGTATGTTGAGAGGACACGAGAGAGAGATTA 1614
DB GGTGCTTCTGTGAGAGGAGACATTTGTATGTTGAGAGGACACGAGAGAGAGATTA 1614
QY 1563 GTGAGCGGAGATGATGAATTAATCTGAGCTACTTGTCTGAAACCGGAAATCTTA 1622
DB GTGAGCGGAGATGATGAATTAATCTGAGCTACTTGTCTGAAACCGGAAATCTTA 1622
QY 1615 CTGAGCAGAGAGATGATGAATTAATCTGAGCTACTTGTCTGAAACCGGAAATCTTA 1674
DB CTGAGCAGAGAGATGATGAATTAATCTGAGCTACTTGTCTGAAACCGGAAATCTTA 1674
QY 1623 AACAGCTGTCTGTGAGGAGCTTATATCTGACGTGAGCAATCTTCAAGCTGTG 1682
DB AACAGCTGTCTGTGAGGAGCTTATATCTGACGTGAGCAATCTTCAAGCTGTG 1682
QY 1675 GAAAGTGTGCTGTGTGAGGAGCTTATATCTGACGTGAGCAATCTTCAAGCTGTG 1734
DB GAAAGTGTGCTGTGTGAGGAGCTTATATCTGACGTGAGCAATCTTCAAGCTGTG 1734
QY 1683 AACATGAGCTGTGAGGAGCTTATATCTGACGTGAGCAATCTTCAAGCTGTG 1741
DB AACATGAGCTGTGAGGAGCTTATATCTGACGTGAGCAATCTTCAAGCTGTG 1741
QY 1735 AGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1794
DB AGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1794
QY 1742 -CCCCCTGATCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1800
DB -CCCCCTGATCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1800
QY 1795 GTCCCTGTATACCCCCCACTTCAAGGAGGCTTCCCAAGTCTCTTNTCCCTTACTCTCC 1854
DB GTCCCTGTATACCCCCCACTTCAAGGAGGCTTCCCAAGTCTCTTNTCCCTTACTCTCC 1854
QY 1801 TCTCTCTCTGTGAGGCTTCTTCTTCTTGTGTGCTGTGAGTGTGCTGTGAGTGTG 1860
DB TCTCTCTCTGTGAGGCTTCTTCTTCTTGTGTGCTGTGAGTGTGCTGTGAGTGTG 1860
QY 1855 TCTCTCTCTGTGAGGCTTCTTCTTCTTGTGTGCTGTGAGTGTGCTGTGAGTGTG 1914
DB TCTCTCTCTGTGAGGCTTCTTCTTCTTGTGTGCTGTGAGTGTGCTGTGAGTGTG 1914
QY 1861 TTTGGTTCCTTCTCTCTCCCAATTTCTCCGCAATCATTAAGCTTCTTCTGAGCTGA 1920
DB TTTGGTTCCTTCTCTCTCCCAATTTCTCCGCAATCATTAAGCTTCTTCTGAGCTGA 1920
QY 1915 TTTGGTTCCTTCTCTCTCCCAATTTCTCCGCAATCATTAAGCTTCTTCTGAGCTGA 1962
DB TTTGGTTCCTTCTCTCTCCCAATTTCTCCGCAATCATTAAGCTTCTTCTGAGCTGA 1962
QY 1921 GCTGTCTTCTATGAGGAGTCTTGTGCAAAACAGCTGTGTGTGATTTTATGACTTA 1980
DB GCTGTCTTCTATGAGGAGTCTTGTGCAAAACAGCTGTGTGTGATTTTATGACTTA 1980
QY 1963 CTGTGTGTGATTTAGCTACTTGTGGAATCAGCTGCTTCTTCTTCTTATGAGTGTG 2022
DB CTGTGTGTGATTTAGCTACTTGTGGAATCAGCTGCTTCTTCTTCTTATGAGTGTG 2022
QY 1981 GGAATGATCTTACAG--AATCTTGTGCAATCAAAAGTGTGATTTGTGGAAGG 2038
DB GGAATGATCTTACAG--AATCTTGTGCAATCAAAAGTGTGATTTGTGGAAGG 2038
QY 2023 AGGAGATTTTATGAGGAGATTTCTTTCACACCAAAACATCTACCGGCTTGTGAAGG 2082
DB AGGAGATTTTATGAGGAGATTTCTTTCACACCAAAACATCTACCGGCTTGTGAAGG 2082
QY 2039 AACAGATTTCTTCT 2052
DB AACAGATTTCTTCT 2052
QY 2083 TACCGATTTCTTCT 2096
DB TACCGATTTCTTCT 2096
```

RESULT 13

ADFS0144

ID ADFS0144 standard; cDNA; 1629 BP.

XX ADFS0144;

XX 12-FEB-2004 (first entry/)

XX

QY 1683 AACATGAGCTCGAGACAGACTCAAGAACCGCGGTGGATTGTTGGACAGACCATT 1742
DB 1561 AACTGACGCTGGAGACAGAACTGAAGAGCAAGAGGTGGATTGTTGCATTGATACATT 1620
QY 1743 CCCC 1746
DB 1621 GTCC 1624

RESULT 14
AB086171
ID AB086171 standard; DNA; 1071 BP.
XX
XX AB086171;
AC 10-SEP-2002 (first entry)
DT
XX
XX Novel human gene. SEQ ID 42.
DE

Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antitumor; antitumor; antitumor;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo-transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.
XX
XX Homo sapiens.
OS
XX WO200250105-A1.
PN
XX
XX 27-JUN-2002.
PD
XX 17-DEC-2001; 2001MO-US049232.
PF
XX 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281353P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PA (GLAX) GLAXO GROUP LTD.
PA
XX
XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KP, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX MPI: 2002-508784/54.
DR P-PSDB: ABP61006.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX Claim 2(a); Page 252; 335pp; English.

The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,

CC cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, neurotropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antitumor, antitumor, antitumor, antitumor, antitumor,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic disease, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder,
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records AB086130-AB086184 represent novel human cDNA's
CC of the invention
XX

SO Sequence 1071 BP; 242 A; 286 C; 254 G; 289 T; 0 U; 0 Other;

Query Match 47.3%; Score 1021.4; DB 6; Length 1071;
Best Local Similarity 99.9%; Pred. No. 4,8e-251;
Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 819 ATACGTCTCCCATGGCCAAAGCTTATTCGCAAGACCATGAGAGTGGGGTGGCC 878
DB 49 ATACGTCTCCCATGGCCAAAGCTTATTCGCAAGACCATGAGAGTGGGGTGGCC 108
QY 879 ATATCCCTTACCTGAGAGCCCATGATTTATGAAAGATGAGAGCTGAGGTGTTGCA 938
DB 109 ATATCCCTTACCTGAGAGCCCATGATTTATGAAAGATGAGAGCTGAGGTGTTGCA 168
QY 939 CATTTCTGTGTACATGCTGAGCTGAGAGCCCTGCTGAGAGTGGTGAAGACAA 998
DB 169 CATTTCTGTGTACATGCTGAGCTGAGAGCCCTGCTGAGAGTGGTGAAGACAA 228
QY 999 CCTCCAGAGAGCTGCTGAGCCCTGAGCCGCAAGAAACATCTTCACTGAGGTGAT 1058
DB 229 CCTCCAGAGAGCTGCTGAGCCCTGAGCCGCAAGAAACATCTTCACTGAGGTGAT 288
QY 1059 GGTGCTTTCTTCTTATGAGCTGATCTGATCTGATCTGATCTGATCTGATCTG 1118
DB 289 GGTGCTTTCTTCTTATGAGCTGATCTGATCTGATCTGATCTGATCTGATCTG 348
QY 1119 CCTTCATCATGAGATCAATTAACACAGAGTGGCTTCTGCTGCTATGAGAGGCT 1178
DB 349 CCTTCATCATGAGATCAATTAACACAGAGTGGCTTCTGCTGCTATGAGAGGCT 408
QY 1179 CCTGAGATCTGAGAGCTGCTGCAAGAGTGGCTTCTGCTGCTATGAGAGGCT 1238
DB 409 CCTGAGATCTGAGAGCTGCTGCAAGAGTGGCTTCTGCTGCTATGAGAGGCT 468
QY 1239 CACATCCGCTCAGATTTGACCTTGTGGCTTAATGAATACCTTCATGACAGACCTCC 1298
DB 469 CACATCCGCTCAGATTTGACCTTGTGGCTTAATGAATACCTTCATGACAGACCTCC 528
QY 1299 CTGACTGAATATCGAGACAGTCTTGTGACCTTGTGAGATGTGTTGTGTGCTCC 1358
DB 529 CTGACTGAATATCGAGACAGTCTTGTGACCTTGTGAGATGTGTTGTGTGCTCC 588
QY 1359 GCACTGATACAGCTCGATATGACAGATGTGAGTGGCACTGTCTACTTCTATGAGTT 1418
DB 589 GCACTGATACAGCTCGATATGACAGATGTGAGTGGCACTGTCTACTTCTATGAGTT 648
QY 1419 CGGACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTGTCAAGCCGACACAGCT 1478
DB 649 CGGACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTGTCTCAAGCCGACACAGCT 708
QY 1479 GATGAAGTCCGTTGTGTGTTGGTGTGCTTCTGAGAGGAGACATGTTATGTTGAA 1538
DB 709 GATGAAGTCCGTTGTGTGTTGGTGTGCTTCTGAGAGGAGACATGTTATGTTGAA 768

Db	707	CCATTATTGCACTGTGGCTGGTGAATACCTTCATGCAAGCACTCCCTGACTGAAA	766
Qy	1309	TCGAGACAGCTCTTCTGACTTGTGAGATGTTCTTTGTGGTCCCTGCACTGATCA	1368
Db	767	TCGAGACAGCTCTTCTGACTTGTGAGATGTTCTTTGTGGTCCCTGCACTGATCA	826
Qy	1369	CAGCTGATATCAAGAGATG	1389
Db	827	CAGCTCAGTATCATAGAGTG	847

Search completed: June 14, 2005, 18:10:04
Job time : 1175.6 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 7160.54 Seconds

(without alignments)
11471.585 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 2158

Sequence: 1 ccacgcgcgcgaacacag.....aaaaaaaaaaaaaaaaaaaaa 2158

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

BST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hnc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gest1:*
9: gb_gest2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	22.2	742	4	BI822069 603039938
2	418.6	19.4	1680	4	AY414461 Mus muscu
3	414.6	19.2	576	4	BI824830 60303738
4	406.8	18.9	435	4	BM422399 HAS05_70
5	406.6	18.8	1680	4	AY414459 Homo sapi
6	397.8	18.4	2661	3	AK033563 Mus muscu
7	396.2	18.4	2687	3	AK077248 Mus muscu
8	387.6	18.0	4927	3	AK040349 Mus muscu
9	385.2	17.8	1958	3	AK078953 Mus muscu
10	353.6	16.5	1944	3	AK078879 Mus muscu
11	345.8	16.0	2888	3	HSW803013
12	344.2	15.9	3909	3	HSW806270
13	338	15.7	1918	7	CF111083
14	333.8	15.5	989	3	AK007235
15	331.4	15.4	2038	3	BC019926 Mus muscu
16	298.4	13.6	1826	3	CR618303 full-len
17	293.6	13.6	1859	3	CR641368 Tetradon
18	292	13.5	1817	3	CR634885
19	289	13.4	1510	3	AK050444 Mus muscu
20	286	13.3	2356	3	AK037191 Mus muscu
21	285	13.2	491	5	BX280446 BX280446
22	284.2	13.2	1849	5	CR647187 Tetradon
23	284	13.2	539	5	BUS85267
24	284	13.2	593	1	AI808985

25	282.2	13.1	805	5	BP442124	BP442124
26	282.2	13.1	808	5	BP446194	BP446194
27	282.2	13.1	840	5	BP445396	BP445396
28	282.2	13.1	841	5	BP447764	BP447764
29	282.2	13.1	860	5	BP441209	BP441209
30	282	13.1	904	5	BQ900930	BQ900930
31	281.8	13.1	798	5	BP444116	BP444116
32	281.8	13.1	808	5	BP443157	BP443157
33	281.8	13.1	877	5	BP441564	BP441564
34	281.8	13.1	877	5	BP442855	BP442855
35	281.6	13.0	1823	3	CR640788	CR640788
36	280.8	13.0	1853	3	CR723405	CR723405
37	280.2	13.0	840	5	BP442111	BP442111
38	279.8	13.0	783	5	BP445505	BP445505
39	279.6	13.0	792	5	BP443197	BP443197
40	276	12.8	645	5	BUS85270	BUS85270
41	275.4	12.8	905	7	CO883389	CO883389
42	274	12.7	1857	3	CR637401	CR637401
43	273.6	12.7	1117	7	CN061747 Ag2_P32_1	CN061747 Ag2_P32_1
44	272.4	12.6	763	7	CR765061	CR765061
45	272	12.6	770	5	BP444251	BP444251

ALIGNMENTS

RESULT 1
LOCUS BI822069 742 bp mRNA linear EST 04-OCT-2001
DEFINITION 603039938P1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181056 5',
mRNA sequence.
ACCESSION BI822069
VERSION BI822069.1 GI:15933619
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 742)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11451 row: k column: 09
High quality sequence stop: 740.
Location/Qualifiers

FEATURES

source

1..742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5181056"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 22.2%; Score 479; DB 4; Length 742;
Best Local Similarity 100.0%; Pred. No. 9.1e-119;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 GGGCTTTGCTGTAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGCGACAA 245
DB 103 GGGCTTTGCTGTAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGCGACAA 162
QY 246 GTCACTGCTGGAAGCCCTGTCCTGTAACGCTGTTCTCGAGTCCCTTTGCTGCT 305
DB 163 GTCACTGCTGGAAGCCCTGTCCTGTAACGCTGTTCTCGAGTCCCTTTGCTGCT 222
QY 306 CCCCCTGGGATCCCTGCAATTTAGCAACCCGAGCTGATGACCTGGAGTAATCTG 365
DB 223 CCCCCTGGGATCCCTGCAATTTAGCAACCCGAGCTGATGACCTGGAGTAATCTG 282
QY 366 CGAAGACCACTCTTACCTTAATTTGCTCTCAAGACTCAGAGTGGCTGCTTAGAT 425
DB 283 CGAAGACCACTCTTACCTTAATTTGCTCTCAAGACTCAGAGTGGCTGCTTAGAT 342
QY 426 CAACATGCTCAAGGTCATTAACCGAATTCGAGTGCAGAGATGCTCTTACCTG 485
DB 343 CAACATGCTCAAGGTCATTAACCGAATTCGAGTGCAGAGATGCTCTTACCTG 402
QY 486 AATATCTATGCGCTGCTCCACGCGATACAGGCTCCAGCTCTTGTGTGTGTC 545
DB 403 AATATCTATGCGCTGCTCCACGCGATACAGGCTCCAGCTCTTGTGTGTGTC 462
QY 546 CCAAGAGTGCCTTCAAGACTGCTCAGGCTCATTTGTAGAGGTCCGCTGCTGCTC 605
DB 463 CCAAGAGTGCCTTCAAGACTGCTCAGGCTCATTTGTAGAGGTCCGCTGCTGCTC 522
QY 606 TATGAGACGTGCTGCTGCTGCTGCTCAATACCGCTAGAAATTTGTTTCTTCA 664
DB 523 TATGAGACGTGCTGCTGCTGCTGCTCAATACCGCTAGAAATTTGTTTCTTCA 581

RESULT 2
AY414461 1680 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus CES2 gene, VIRTUAL TRANSCRIPT, partial sequence.
DEFINITION AY414461
ACCESSION AY414461 GI:39770423
VERSION GSS.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1680)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Clavello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Interfering nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL PUBMED 14671302
REFERENCE 2 (bases 1 to 1680)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Clavello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
SOURCE 1..1680
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

Gene <1..>1680
/gene="CES2"
/locus_tag="HCM5231"

Query Match 19.4%; Score 418.6; DB 9; Length 1680;
Best Local Similarity 56.9%; Pred. No. 3.5e-102;
Matches 885; Conservative 0; Mismatches 634; Indels 36; Gaps 5;

ORIGIN

QY 210 AGGAACACAGGCTGGATGATTCAGGGCAAGCAAGTCACTGCTGGAGGCTGTG 269
DB 100 AGGAACACACAGGCTGGATGATTCAGGGCAAGCAAGTCACTGCTGGAGGCTGTG 159
QY 270 CCTGTAACGTTGCTCTGAGATCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 329
DB 160 GGTGTGCACTGCTCTGAGGATCCCTTTGCTGAGCTCTGTAAGCACTGCTGCTT 219
QY 330 ACGAACCAGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389
DB 220 GCAACCCCTGAGCACTGAAACATGAGGAGTGTGTGAGAGATGGAGATCATCATCGGGCC 279
QY 390 TTGTGCTCCAG---AACTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 446
DB 280 ATGTGTGCAAGACATCATCTGCAATGATGAGGCTTTTAAGCTGTTGAAGCTGACC 339
QY 447 TACCGAAATTCGAGTGTCAAGAACTGCTTACCTGAAATCTATGCGCTGCTGCTC 506
DB 340 TTGCTCTCTCTCTATATGTCTGAGACTGCTCTTAACTTAATGATGACCAATCAT 399
QY 507 GCGATACAGGCTTCAAGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 566
DB 400 GCCATAGAGGTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459
QY 567 GCTCAAGCTTCACTTTGATGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 626
DB 460 GCGATGCTTCAATGATATGATGATCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
QY 627 GTGCTCAAGTACCGGCTGAGAAATTTGCTTCTTCAACATGAGATGAGATGCTGCTG 686
DB 520 ACTATCAGATACCGGCTGAGTGTCTGCTGCTTCTTCAAGCATGAGGATCCAGATCCAGA 579
QY 687 GGGAACTGAGGCTTCAAGAACAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
DB 580 GGCATTTGGGATACCTGATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639
QY 747 TTTTGTGTTGGGAGCCCACTGCTGCAATCTTTGGGAGTCCGCGAGCCATAAGT 806
DB 640 CACTTTGTGAGCAACCGTGAACGGGTTACAAATTTTGGGAGTCTGCAAGTGGCAAGT 699
QY 807 GTTCTAGCTTATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 866
DB 700 GTCTCTTCACTGTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
QY 867 AGTGGGTTGGCATCATCTTACCTGAGAGGCCATGATTTAGAAAGATGAGAGACTG 926
DB 760 AGTGGGTTGGCATCTGCTGCTGCTTACTTATCTCAACACTCTGAGA-----TGTCTCC 813
QY 927 CAGGTGTTGCACTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 986
DB 814 ACTTACGTTGGCCAGACTCTGAGATGAGGCGGTGACTGAGAGGCTGCTGCTGCTGCT 873
QY 987 CTGAGACAAACCTCCAGAGAGCTGCTGACCTGAGCCGAGAAACAAAGTCTTTCACT 1046
DB 874 CTGAGAGCAAGATGAGAGCAAGATTTCTGCTTATTAAGAACTTCCAGATATCTCT 933
QY 1047 CGAGTGTGATGCTGCTTTTCTTATAGAGCTTATGATCTATGCTCAGAAAGCA 1106
DB 934 GGTGTGATGATGAGATGCTTCTACCAAGCATCTCAGAGAGCTGCTGCTGCTGCTGAT 993
QY 1107 TTAAAGCAATTCCTCATCATGAGTCAATTAACAGAGTGTGCTTCTGCTGCT 1166
DB 994 TTTCACCTGTGCTCCAGCATCATTTGTTTCAACATGATGATGATGATGATGATGATGAT 1053

OY	1167	-----ATGAAAGAGGCTCTGAGATCTCAATGGCTCCAAACAAGCTTC	121.4
Db	1054	AAAAGCATGAGCATATGATCTCCAGAAGGGATGACACAGCAATGCTGACACTC	111.3
OY	1215	CTCCATCTGATACAAACATCCTGCATCCCGCTCAGTATTGCACTTGTGGCTAT	127.4
Db	1114	CTGCAGAGAGGGCAACCCAGATGATGTGCTCTTGATGTAGTACCTCTATGGA	117.3
OY	1275	GAATATCTTCATGACAGACATCCCTGACTGAAATCCAGACAGTCTTGTGACTT	133.4
Db	1174	GAGTACAGAGGGGACACTGAGAGTCCAGACCCCTCAAAATACAGTTCACAGATGAT	123.3
OY	1335	GGAATGTGTTCTTTGGTCCCTGCACTGATACAGTGCATATCAAGAGATGCTGT	139.4
Db	1234	GGAATTTTCAATGTTTGTATCCCTGCACTCCAGTAGACATTTCAAG--TTCCAT	129.0
OY	1395	GCACCTGTCTACTTATAGTTTCGACACCGGCTCAGTCTTTGAAGACAGAACCG	145.4
Db	1291	GCCCTGTCTACTTCTATGAAATCCAAATCAATCACTTCCTCCAGAGATGACGGCA	135.6
OY	1455	GCTTTGTCAAAGCCGACACGCTGATGAAGTCCGCTTGTGTGGTGGTCTTCTG	151.4
Db	1351	CCCCACGTGAACGCTGACCATGAGATGAGTCTTTGTGTTTGGTCTCTTCTTCGG	141.0
OY	1515	AAGGGGACATTTGTTATGTTGGAAGGACCCAGAGAGAGGAAGTTACTAGCCGAG	157.4
Db	1411	GGCATGAACCTTGACTT-----CACTGAGGAGAGAGCTGTAGACAGAGG	145.8
OY	1575	ATGATGAATATCTGGGCTACCTTTGCTGAAACCGGGAATCCTTAATGGGAACGACTGT	163.4
Db	1459	ATGATGAAGTACTGGGCCCATTTTGCAAGCATGGGAACCCCAACAGTAGAGGTCTAC	151.8
OY	1635	CTGTGGCCAGCTTATATCTGACTGAGCAATCTCCAGCTGGAAGTTGAACATGAGCTC	169.4
Db	1519	TACTGGCTGTGATGAGACCATGATGAGCACTGCAAGCTGGAATCAACCTTGTGTG	157.8
OY	1695	GGACAGAGACTCAAGAACCGCGGGTGATTTTGGACACACACATCCCGCTA	174.9
Db	1579	GGTGAGGCTCGAAGGCCAGAAAGCTGCAAGTTCTTGACCAAGACTCTGCCCA	163.3

[illegible]

```

FEATURES
    source
        CDNA library preparation: Life Technologies, Inc.
        cDNA library arrayed by: The I.M.A.G.E. Consortium (LNL)
        DNA sequencing by: Incyte Genomic, Inc.
        Clone distribution: MGC clone distribution information can be
        found through the I.M.A.G.E. Consortium/LNL at:
        http://image.llnl.gov
        Plate: LHAM1435 row: 1 column: 10
        High quality sequence start: 6
        High quality sequence stop: 565.
        location/Qualifiers
            1..576
            /organism="Homo sapiens"
            /mol_type="mRNA"

```

```

/db xref="taxon:9606"
/clone="IMAGE:5174937"
/lab host="DH10B"
/clone_lib="NH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site_1: NotCl; Site_2: ECoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (ECoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NH_MGC library."

ORIGIN

Query Match      19.2%; Score 414.6; DB 4; Length 576;
Match Local Similarity 96.4%; Pred. No. 3e-101;
Matches 456; Conservative 0; Mismatches 14; Indels 3; Gaps 3;
```

Query Match	Similarity	19.2%;	Score 414.6;	DB 4;	Length 576;	
Best Local	Similarity	96.4%;	Pred. No. 36-101;			
Matches	456;	Conservative	0;	Mismatches 14;	Indels 3;	Gaps 3;
Qy	186	GGGCGTTCTGCTGAAAGGCGCA	CAGAGNACACAGCGCTGGATGATTCAGGGCAGCA	245		
Db	104	GGGCGTTCTGCTGAAGGGCCAC	AGAGNACACAGCGCTGGATGATTCAGGGCAGCA	163		
Qy	246	GTCACTGTGCTGGAAAGCCCTG	GCTGTGAACGTGTTCTTCGAGTCCCTT-TGCTGC	304		
Db	164	GTCACTGTGCTGGAAAGCCCTG	GCTGTGAACGTGTTCTTCGAGTCCCTTATGCTGC	223		
Qy	305	TCCCCCGCTGGGAATCCCTTG	CGATTTAGAAACCGGACGCTGCATCCGCTGGGATTAATT	364		
Db	224	TCCCCCGCTGGGAATCCCTTG	CGATTTAGAAACCGGACGCTGCATCCGCTGGGATTAATT	283		
Qy	365	GCGAAGAGCCACTCTCAACCT	TAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTTAGA	424		
Db	284	GCGAAGAGCCACTCTCAACCT	TAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTTAGA	343		
Qy	425	TCAACACATGCTCAAGGTGCA	TTTACCCGAAATTCG-AGTGTCAAGAACCTGCTTACC	483		
Db	344	TCAACACATGCTCAAGGTGCA	TTTACCCGAAATTCGAGTGTCAAGAACCTGCTTACC	403		
Qy	484	TGAACATCTATGCGCTGCGCC	ACGCCAGATACAGGCTCCAGCTCCCGCTTTGTGTGT	543		
Db	404	TGAACATCTATGCGCTGCGCC	ACGCCAGATACAGGCTCCAGCTCCCGCTTTGTGTGT	463		
Qy	544	TCCAGAGAGGTGCTTTCAAG	ATGAGTCAAGCTCATCTTTGATGGGTGGCCGCTGGGTG	603		
Db	464	TCCAGAGAGGTGCTTTCAAG	ATGAGTCAAGCTCATCTTTGATGGGTGGGTGGCTGGGTG	522		
Qy	604	CCTATGAGGACGTGCTGTTGT	GTGCTGCTCCAGTACCGGCTAGAAATATTTGTT	656		
Db	523	CTATGACGAGCTGTGGTGTGT	GTGCTGCTCCAGTACCGGCTAGAAATATTTGCT	575		

RESULT	4
BM422399	
LOCUS	
DEFINITION	BM422399 Homo sapiens testis CDNA library Homo sapiens CDNA, mRNA
ACCESSION	
VERSION	BM422399
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 435) Liang,G., Mao,S.Y., Zhang,X.D., Wang,L.J. and Wang,L.F. The differentially expressed gene in spermatogenesis Unpublished (2002)
REFERENCE	
AUTHORS	Contact: Lin-Fang Wang
TITLE	National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences & Peking Union Medical College; Chinese National Human
JOURNAL	
COMMENT	

Genome Center, Beijing
#5, Dong Dan 3 Rao, Beijing 100005, P.R.China; #3-707 North
Yongchang Road BDA, Beijing 100176, P.R.China
Tel: 8610-65396418
Fax: 8610-65240529
Email: wangli@cdm.lmicams.ac.cn.

FEATURES
Source
Location/Qualifiers

1..435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="male"
/issue_type="testis"
/dev_stage="adult"
/clone_lib="Human testis cDNA library"
/note="The cells of Primary Spermatoocyte and Round
Spermated were isolated from Human testis. Total RNA was
extracted by Trizol. Using PCR-Select cDNA Subtraction
Kit, the library was constructed"

ORIGIN

Query Match 18.9%; Score 406.8; DB 4; Length 435;
Best Local Similarity 99.5%; Pred. No. 3.7e-99;
Matches 408; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1666 ACCCTCCAGCTGGAAGCTGACATGAGCTCGACAGAGACTCAAGAGCGGGTGGATT 1725
DB 1 ACCCTCAGCTGGAAGCTGACATGAGCTCGACAGAGACTCAAGAGCGGGTGGATT 60
QY 1726 TTGGAACGACAGCAATCCCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 1785
DB 61 TTGGAACGACAGCAATCCCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 120
QY 1786 CTCTCTTAATCT 1845
DB 121 CTCTCTTAATCT 180
QY 1846 TTATCTTCTGATTTTGGTTTCCCTTCTCTCTCCATTAATTTCTCCGCAATCTAGC 1905
DB 181 TTATCTTCTGATTTTGGTTTCCCTTCTCTCTCCATTAATTTCTCCGCAATCTAGC 240
QY 1906 TTCTTTCTGAGCTGAGCTGCTTTCTATGGGGATCCCTGCAAAAGAGCTTGGCTGA 1965
DB 241 TTCTTTCTGAGCTGAGCTGCTTTCTATGGGGATCCCTGCAAAAGAGCTTGGCTGA 300
QY 1966 TATTTTGAAGATTGGAATGATCTTCAAGAAATCTTTTCAACATCAAAAGTGCATT 2025
DB 301 TATTTTGAAGATTGGAATGATCTTCAAGAAATCTTTTCAACATCAAAAGTGCATT 360
QY 2026 TGTCTTGAAGGCAAGAAATTTCTTCAATAATTTGGAAGAGGCTGGC 2075
DB 361 TGTCTTGAAGGCAAGAAATTTCTTCAATAATTTGGAAGAGGCTGGC 410

RESULT 5
LOCUS AY14459 1680 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens CES2 gene, VIRUAL TRANSCRIPT, partial sequence,
ACCESSION AY14459
VERSION AY14459.1 GI:39770421
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
gene trios
Inferring nonneutral evolution from human-chimp-mouse orthologous

JOURNAL
PUBMED
REFERENCE
AUTHORS
14671302
2 (bases 1 to 1680)

Clark,A.G., Glanowski,S., Nelson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES

source
Location/Qualifiers
1..1680
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/gene="CES2"
/locus_tag="HCM5231"

ORIGIN

Query Match 18.8%; Score 406.6; DB 9; Length 1680;
Best Local Similarity 56.5%; Pred. No. 6.5e-99;
Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;
QY 211 GGAACACGAGCTGGATGATTCAGGGCAAGCAATCACTGTGTGGAAAGCCGTGCG 270
DB 101 GGAACACACGAGCTGGATGATTCAGGGCAAGCAATCACTGTGTGGAAAGCCGTGCG 160
QY 271 CTGTGAACGATTCCTGCGAGTCCCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 330
DB 161 GGTTCAAACCTTCCTGGAATTCATTTGCCAAGCACTTCTAGGTCGCTGCAATTG 220
QY 331 GGAACCCGACCTGATGCGCTCGGATTAATTTGCAAGAGCACTCTTACCTTAAT 390
DB 221 CACCCCTGAGCCCTCGAATCTTGAAGTGTGTAGAGATGAAACCACTTCGGCCA 280
QY 391 TGTGCTCCGAACTAGATGAGTCTCTTGA--TCMAACAATGCTCAAGTGCATT 447
DB 281 TGTGCTCCGAACTAGATGAGTCTCTTGA--TCMAACAATGCTCAAGTGCATT 340
QY 448 ACCGAAATTCGAGATGATGAGAGATGCTCTTACCTGAACATTAATGCGCTGCGCAC 507
DB 341 TCCCTTCGACATTCATGATGATGAGATGCTCTTACCTGAACATTAATGCGCTGCGCAC 400
QY 508 CCGATACAGGCTCCAACTCTCCCGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 567
DB 401 GCGATGAAAGGCTTAACCTGCGGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 460
QY 568 GCTCAGCTCCATCTTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 627
DB 461 GCTCAGCTCCATCTTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 520
QY 628 TGTTCAGTACCGGCTAGAAATTTGTTTCTTACCACTGAGATCAGACTCCGG 687
DB 521 TCATCAGATACCGGCTAGGATGCTGAGCTTCTTACCACTGAGATCAGACTCCGG 580
QY 688 GGAACCTGGGCTTCAAGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 747
DB 581 GGAACCTGGGCTTCAAGACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 640
QY 748 TGTTCAGTACCGGCTAGAAATTTGTTTCTTACCACTGAGATCAGACTCCGG 807
DB 641 ACTTGAAGGAAACCTGATACCTGTACATTTTGTGAGATGCTGAGATGCTGAGATGCT 700
QY 808 TTTCTAGCTTATATCTGTCTTCCATGAGCAAGCTTAATTCACAAAGCATATGAGA 867
DB 701 TGTCTTGTGCTTGT 760
QY 868 GTGGGATGAGCATATCTTACCTGAGGCGCATATTTGAGAAAGTGTGAGAGCTGCTG 927
DB 761 GTGGGATGAGCATATCTTACCTGAGGCGCATATTTGAGAAAGTGTGAGAGCTGCTG 814

Oy 928 AGTGGTTGCACTTTCTGTGTAAACAATGCGTCAGACTGAGGCCCTGCTGAGTGGC 987
 Db 815 CGGTGTGTCACCACTGTCTGTCTGTACCAAGTTGACTCTGAGGCCCTGTGTGGCTGCC 874
 Oy 988 TGAGGACAAAACCTCCCAAGAGCTGTGACCTCAGCCAGAAAACAAAGTCTTTCCTC 1047
 Db 875 TGCCGGGCAAGATGATTAAGAGGATTTCTTGCAATTAAAGACCTTCAAGATGATCCCG 934
 Oy 1048 GAGTGTGTGATGAGTCTTCTTCTTCTTAATGAGCCCTGAGTCTATTTGTCTAGAAAGCAT 1107
 Db 935 GAGTGTGTGATGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 994
 Oy 1108 TTAAGCAATTCCTCCATCATCGAGTCAATTAACAGAGTGTGCTTCTGCTGCTCA 1167
 Db 995 TTCAAGCTGTCTCTGATTTGTTGTGTCAACAATGAAATTCGGCTGGCTCAATCCCA 1054
 Oy 1168 TGA-----AGAGGCTCTGAGATCTCAGTGGCTCAACAAGTCCCTTGGCC 1215
 Db 1055 AGGTGATGAGATCTATGATTCAGAAAGAAATGAGACAGAGGCTCCAGCGTGC 1114
 Oy 1216 TCCATCTGATACAAAACATCTGACATCCGCTCAGTATTTGCACTTGTGCTAATG 1275
 Db 1115 TGCAAAAATGTAAAGCTGATGTTGCTCTTCACTTGTGTGATGCTGCTGAGGAGG 1174
 Oy 1276 AATATCTTCATGACAGACATCCCTGAGTGAATCGAGACAGTCTTGTGACCTTGTG 1335
 Db 1175 AGTACATTTGGGACATAGGGAGATCCCAAGACCTCTCAAGCCGATTCAGAGATGATG 1234
 Oy 1336 GAGATGTTCTTTGTGTGCTCCCTGACATGATCAGCTGATCAGATCAGAGATGCTGTG 1395
 Db 1235 CGGACTCATGTTTGTGTGATCTCCGACCTCAGATGACATTTTCAGTG---TTCCCGG 1291
 Oy 1396 CACCTGTCTCTTCTATGATGTTGGGACCGGCTCAGTGTCTTGAAGACAGAACCGG 1455
 Db 1292 CCCTGTCTCTTCTTCAAGATTCAGATCCAGACCGGCTCAGAACATCAGGCCAC 1351
 Oy 1456 CTTTGTCAAGCCCAACAGCTGATGAAGTCCGCTTGTGTGCTGTGCTCTTCTGA 1515
 Db 1352 CGCATATAGGCAACACATGTGTGATGATCTTCTTGTGTTTCAAGATTTCTTGGGG 1411
 Oy 1516 AGGGGACATTTGTTTTCAGAGAGCCAGAGAGAGAGATTAAGTCAAGCCGAGAA 1575
 Db 1412 GCACTACATTTAAT-----CAGTGAAGAGAGAGAGAGAGAGAGAGAA 1459
 Oy 1576 TGATGAATTAATCGGCTACCTTGTCTCAACCGGAAATCTTAATGGAACGACTGTCTC 1635
 Db 1460 TGATGAATTAATCGGCTACCTTGTCTCAACCGGAAATGGAACCGGAGGCTGTGCAC 1519
 Oy 1636 TGTGGCCAGCTTAATATCTGATGAGTGAAGTCTCAGCTGATCTTGAACATGAGCCCTG 1695
 Db 1520 ACTGGCCGCTGTTCAGACAGAGAGAGCAATATCTCAGCTGAACCTACGCTGGGTGG 1579
 Oy 1696 GACAGAGACTCAAGAACCGGCGGTGATTTTGGACCGACCATCCCC 1746
 Db 1580 GCCGGGCTCTGAAGGCCCAAGGCTCCAGTTCTGAAGAAGCGCTGCCCC 1630

RESULT 6
 AK033563
 LOCUS
 DEFINITION
 AK033563 2661 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030624L02 product:similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1) (AL-ESTERASE) (B-ESTERASE) (MONOESTERASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLESTERASE) (mesocricetus auratus), full insert sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AK033563.1 GI:26329244
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 PUBMED
 99279253
 10349636
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 PUBMED
 20493374
 11042159

TITLE

The RIKEN Genome Exploration Research Group Phase II Team and the
 PANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

TITLE

The PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

REFERENCE

6 (bases 1 to 2661)
 Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

TITLE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-research@riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in RIKEN
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in RIKEN contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/.
 Location/Qualifiers

FEATURES

1..2661
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM_DB:9030624L02"

```

/db xref="taxon:10090"
/clone="9030624L02"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
25..11704
/notes="unnamed protein product; putative
similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1)
(ALT-ESTERASE) (B-ESTERASE) (MONOESTERASE) (COCAINE
ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
[Meocricetus auratus] (SPR/035533, evidence: FASTY,
83%ID, 100%length, match=1677)"
/codon_start=1
/protein_id="BAC28361.1"
/db xref="gi:26329245"
/translation="MPYKLLGLMNAVACGVLLLVHVGQDSASPIRHTHTGVRGS
LVYKQDIDIAVHTFLGIPFAKPPVGLRFAPEAPPEMVGSRDGTSHPMCLQNDMLM
GSBLKMNLLLPISMEDLYLNIYVPAHAGSNLPVWVHGGALTVGASMYD
GSMILATEDVVVAIQRGLVGFSTGDHAKGNMGYLDQVLAALMVQONTVHREGN
PDRVTLFESAGTSVSHVSPMSQGLFGLAIMESGVATLPOLISSSEMHRIYAN
LSGCAVNSETLMCCLEKNEAEMLAIINKFKIIRPVVSEPIPKRPOELMASKDFP
VPSIIGNDEYGMILPTIMDPAOKIETIRKTLPAVLSTALKMLPECGDLMEF
YMGPTDEPRLQAFPMKCDPMFVIPALQVAHFQSHAPVYVFEQHRSPFKQFRP
PYKADHDEITFVFEQIFQSNITLPTVEEEOQLSRMXYRANFAHGNPSEGLPYW
PVVDHBOYLQDIIQPSVGRALARRLQFWTKTLPKIOLKSGQRHKEI"
2634..2639
/notes="putative"
2661
/notes="putative"

```

```

polya_signal
polya_site
ORIGIN

```

```

Query Match      18.4%; Score 397.8; DB 3; Length 2661;
Best Local Similarity 56.1%; Pired. No. 1.9e-96;
Matches 872; Conservative 0; Mismatches 647; Indels 36; Gaps 5;

210 AGGAAACAGGCTGGATGATTCAGGGCAGCAAGTCACTGTCTGGAAACCCCTGTG 269
    |||
Db      124 AGGAAACAGCAACAGGCGAGTGAAGGACGCTTGTCCAGTAAAGACCTGACATC 183

QY      270 CCTGTACGCTGTCCTCGAAGTCCCTTGTGCTGCCCGCTGGAGTCCCTGGATTT 329
    |||
Db      184 GGTGTTCATACCTTCTCGGAATCCCTTGTGCAAGCCCTCGTAGACCACTGGGCTTT 243

QY      330 AGGAAACCGGACCTGCAATCGGCTGGGATTAATTGCGAAGACCACTCTCAACCTAAT 389
    |||
Db      244 GCACTCTCGAGGCCCTTGAGCCATGAGTGGGAGTGGAGATGGCAATCAATCAAAAC 303

QY      390 TTGTGCTTCAGAACTCAGAGTGGCTGCTTAA--GATCAACACATGCTCAAGTGCAT 446
    |||
Db      304 ATGTGCTACAAATGACAAATTGATGGGTTCAAGAGATCTGMAATGATTAATCTATTC 363

QY      447 TAACCGGAATTCGAGAGTCTGAGAAAGTCTCTAATCTGAACATCTATGCGCCCTGCCAC 506
    |||
Db      364 CTGCTCCCATCTCTATGTCTGAGGACTGCTGTATCTCAACATCTATGTATCAAGGCCAT 423

QY      507 GCCGATACAGGCTCCAAGCTCCCGCTCTGGTGTGTTCCAGAGAGTGGCTTCAAGACT 566
    |||
Db      424 GCCCATGAGGGTTCTAATCTGATGATGATGATGATCAATGATGATGATGATGATGATGAT 483

QY      567 GGGTCAGCTCTCAATCTTGAATGGTCCGCTGGCTGCTATGAGAGAGTGGTGTGTTG 626
    |||
Db      484 GGGATGGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543

QY      627 GTCTGCAATACCGGCTAGGAATATTGTTTCTTCAACCAATGGAATCAGATGCTCCG 686
    |||
Db      544 GCTATCAAGTACCGGCTGGGTTCTTGAATCTTCAACCAATGGAATCAGATGCTCCG 603

QY      687 GGGAACTGGGCTTCAAGACCAAGTGGCTGTCTGTCTGGGTCCAGAAACATTCAG 746
    |||
Db      604 GGGAACTGGGGAATACCTGACCAAGTGGCTGTCTGTCTGGGTCCAGAAACATTCAG 663

QY      747 TTCTTGGTGGGAGCCCAAGCTCTGTGACATCTTTGGGAGTCCGCGGAGGACATTAAGT 806

```

```

Db      664 CACTTGGAGGACACCTTCAACGAGTCACTTTTGGAGTCAAGCGGGGCACAAGT 723
QY      807 GTTTCATGCTTAACTGTCTCCATGGCCAAAGGCTTATTCACAAAGCATCATGGAG 866
Db      724 GTGTCTTCAATGTGTGTGTCCTCCCATGTCCTCCAGGAGCTTTCATGTGTCATATGAG 783
QY      867 AGTGGGTGGGACCATCATCCCTTACCTGAGAGCCCATGATTTATGAGAAAGTGAAGACCTG 926
Db      784 AGTGAAGTACGTGTCTCTCTGATCTTATCTCCAGCTTCTTGAGA-----TGGTTAC 837
QY      927 CAGGTGTTGACATTTCTGTGTAACTGCTCAGACTGTGAGGCGCTGTAGAGTGC 986
Db      838 AGAATAGTGGCCAACTCATCTGTGGCTGACAGCTGTGAATCAAGACCTGTATGTGCTGC 897
QY      987 CTGAGAGCAAAACCTCCAAAGAGCTGTGACCTCAGCCAGCAAAACAAAGTCTTCACT 1046
Db      898 CTAAAGAGGCAAGAAATGAAGCAGAGATCTGTCTTAAATAGGCTTCAAAATCATCCCT 957
QY      1047 CGAGTGTGTTAGTGGTCTTCTTCTTCTTATGAGCTCTAGATCTATGTCTCAAGAAAGCA 1106
Db      958 GGTGTGATGATGAGAGTTCCTTACCAAGACTCTCAGAGCTGATGTGCTTAAAGAT 1017
QY      1107 TTTAAAGCAATCTCTTCAATCATGAGTCAATPAACACAGTGTGCTTCTGTGCTCT 1166
Db      1018 TTGACCCCTGCTCCAGCATCATTTGTATCAACATGATGAATATGTTGATTTCTTCT 1077
QY      1167 ATGAAGAGGCTCTCT-----GAGATCTCAGTGTGCTCCCAACAAGTCCCTTGC 1214
Db      1078 ACGATCATGAGCCAGCTCAGAAATATGAGAAATTAACGAAGACCTGCGACGCTGT 1137
QY      1215 CTCATCTGATTAACAAACATCTGCAATCCGCTCAGTATTTGCACTTGTGCTAAT 1274
Db      1138 CTAAAGAGCACACCTTAATAATGATGTGCTCTGAGTGTGTGACCTGTATAGAG 1197
QY      1275 GAATACTTCCATGACAGACCTCCCTGACGTAATCCGAGACAGTCTTGTGACTTGTCT 1334
Db      1198 GAGTACATGGGGGACACTGAGAACCCAGAGACTCTGCAAGCAACAGTTCAAGATGAAAG 1257
QY      1335 GAGATGTGTTCTTGTGTGCTCCCTGACATGATCAAGCTCGATATCAAGAGATCTGGT 1394
Db      1258 GGGAGCTTCATGTTCGTGATCCCTGACATCAAGTATGACATTTTCA--ACGTTCCAT 1314
QY      1395 GCACGTGCTAATCTTATGATGTTTGGGACCGGCTCAGTGTCTTGAAGACAGAGCCG 1454
Db      1315 GCTCTGTCTAATCTTATGATGTTTCAACATGACACCACTTCTCAAGATTTAGGCGCA 1374
QY      1455 GCTTTTGTCAAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCGTGTGCTCTCTG 1514
Db      1375 CCTATGTGAAGGCTGACCAATGTATGAATTTTCTTGTCTTGTGCTTCAAGTTCGGT 1434
QY      1515 AAGGGGACATTTGTATGTTGAAAGAGCCACGAGAGAGAAATTATCTGAGCCGGAAG 1574
Db      1435 AACATTAACCTTCTTAA-----CACTAGAGAAAGAGAGCAATTAAGCAAGAG 1482
QY      1575 ATGATGAATAATCGGGCTAACCTTGTGCTGAACCGGGAATCTTATGGAAGACACTGTCT 1634
Db      1483 ATGATGAAGTACCGGCTCACTTTGCAAGGACAGGAAACCCCAACATGAGGCTTACCC 1542
QY      1635 CTGTGGCCAGCTTATATCTGACTGACAGTAACTCTCAGCTGACTTGAACATGAGCTTC 1694
Db      1543 TACTGGCTCTGATGAGACCATGATGACAGATTACTGTGAGCTGAGCATCCAGCTTCTGTG 1602
QY      1695 GGAACAGACTCAAAAGAACCGCGGGTGGATTTTGTGACACGACACATCCCTCTGA 1749
Db      1603 GGTGAGGCTTGAAGGCCAAGAGCTCAAGTTCTGAGCAAGACTGTGCCCAAGA 1657

```

```

RESULT 7
AK077248
LOCUS
DEFINITION
AK077248
Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
RIKEN full-length enriched library, clone:503145B19
2687 bp mRNA linear HTC 03-APR-2004

```


D	512	GGATAGGCTTCATGTATATGATGATTCATGCTGGAGCACTGAGAGTGTGTGTGTC	571
OY	627	GTCTGTCAGATACCGGCTAGAGAAATATTTGGTTCTTCAACAATGGGAATCAGATGCTCCG	686
D	572	GCATATCAATACCGCTGGGTGTCTGGGATTTCTTACACATCGAGACACACATGCCCCAA	6311
OY	687	GGGAATCTGGACCTTCAAGGACCAAGTGTGTCTGTCTGTGGTTCAGAAACATTCGAG	746
D	632	GGCAATGGGGATACCTGGACCAAGTGTGTGCTTCCGTTGGGTTCAGCAAAACATGTGC	691
OY	747	TTCTTGGTGGGAGCCCCAGCTCTGTGACCAATCTTTGGCGAGTCCGCGGGAGCATTAAT	806
D	692	CACTTTGGAGGCAACCTTGACCGAGTCAACATTTTGGAGGTACAGAGGTGGCAATGT	751
OY	807	GTTTCTAGCTTATACGTCTCTCCATGAGCCAAAGCTTATTCACAAAGCATCATGAG	866
D	752	GTGTCTTCAATGTGTGTCTCCCATGTCCAGAGGACCTTCCATGGTGCATCATGAG	811
OY	867	AGTGGAGTGGCCATCATCTTACCTGGAGGCCCATGATTTATGGAAGATGAGACCTG	926
D	812	AGTGAAGTAGCTGTGTCTCTGTATCTTATCTCCAGCTCTCTGAGA-----TGTTCAC	865
OY	927	CAGGTGTGTGACATTTCTGTGTAAACAATGCTCAGACTCTGAGGCCCTGTGAGTGC	986
D	866	AGATATGTGSCAAATCTATCTGTGTGTCAGCTGTGAACTCAAGACCTGTATGTCTGC	925
OY	987	CTGAGACAAACCCCTCAAGAGCTCTGACCCCTGACCGAGAAACAAAGCTTTCACT	1046
D	926	CTAAGAGGCAAGATGAGAGAGAGTCTGGCTATTATTAAGTCTTCAAAATCAATCCT	985
OY	1047	CGATGTGTATAGTGTCTTTCTTCTTAATGAGCCTCTAGATCTATTTGTCTGAGAAACA	1106
D	986	GGTGTGTGTGATGAGATTCACCCAGACATCTCAAGAGCTGATGCTCTAAGAT	1045
OY	1107	TTTAAAGCAATCTCTTCATCATCTGAGATCAATPAACAAGATGTGCTCTGTCTGCT	1166
D	1046	TTTCAACCTGTCCCGACGATCATTTGATATCAAGATGAATATGTGTGATTTCTCT	1105
OY	1167	ATGAGAGGCTCT-----GAGATCTCAGTGTCTCCAAAGTCCCTTGC	1214
D	1106	ACGATCATGAGACCCAGCTCAGAAATATGAGAAATPAACAAGAAACCTGTGCAGCTGT	1165
OY	1215	CTCCATCTGTATCAAAACAATCCTGCAATCCGCGCTCAGTATTTGACACTTGTGGCTAT	1274
D	1166	CTTAAAAAGCAGCCCTTAAAAATGATGTCTGCTCTGATGTGTGATCTCTATAGAG	1225
OY	1275	GATACTCTTATGACAAAGCACTCTCCCTGACTGAAATCGAGACAGTCTTGTGACTTCT	1334
D	1226	GAGTATATGGGGAGCACTGAGACCCAGAGACTGTGCAAGCAGTTCAAGAGATTAAG	1285
OY	1335	GGAGATGTGTCTTGTGTGTCTCCCTGACATGATCAAGTCCGATATCAACAGATGTGTGT	1394
D	1286	GGGAGCTTATGTGTGTATCTCTGTATCCCTGTACTCAAGTATGACATTTTCA---ACGTTCCAT	1342
OY	1395	GCACTGTCTACTTCTATGATGTTTGGGACCGGCTCAGTGTCTTGTGAAGAACAAGACCG	1454
D	1343	GTCTCTGTCTACTTCTATGATGTTCAATCATGACCCAGCTTCTTCAAGATTTCAAGCCA	1402
OY	1455	GCTTTGTCAAGCGGACACGCTGATGAATCCGCTTTGTGTCTGGTGTGCTTCTGT	1514
D	1403	CCCTATGTGAAGGCTGACCAATGTGTATTAATTTCTTGTGTGTGCTACAGTTTCGT	1462
OY	1515	AAAGGGGACATTTGTATGTTGTGAAGAGCCAGAGAGGAGAAATTACTGAGCCGGAAG	1574
D	1463	AAACATAAATCTCTTA-----CACTAGAGAAAGAGAGCACTTACAGAAAG	1510
OY	1575	ATGATGAATACTGGGCTACCTTGTCTGAAACGGGAAATCTTAATGGAAACGACTGTCT	1634
D	1511	ATATATGAATGTAATGTGGCAACTTTTGACCGGACGGGAACCCCAACAGTGTGGGTCTACCC	1570
OY	1635	CTGTGGCCAGCTTAAATCTGATGACAGATTAACCTCAGCTGAGACTTGAACATGAGCTTC	1694
D	1571	TACTGTGCTGTGATGAGCAATGATGACAGATTAACCTGAGCTGAGCAATCCAGCTTCTGTG	1630

QY	1695	GACAGAGATCTCAAGAACC	CGGGTGTGATT	TTTTTGACACAGACCATTC	CCCTGA	1749
Dd	1631	GGTGCAGCCTTGAAAGCC	CAGAAAGGCTG	CTGTGACCAAGACTCTG	CCCCA	1685
RESULT 8						
LOCUS	AK040349					
DEFINITION	Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A43008B12 product:similar to CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) (AL-ESTERASE) (B-ESTERASE) (MONOBUZYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE) [Mesocricetus auratus], full insert sequence.					
ACCESSION	AK040349					
VERSION	AK040349.1	GI:26087790				
KEYWORDS	HTC; CAP trapper.					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	Carninci, P. and Hayashizaki, Y.					
TITLE	High-efficiency full-length cDNA cloning					
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)					
MEDLINE	99279253					
PUBMED	10349636					
REFERENCE	2					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.					
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes					
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)					
MEDLINE	20499374					
PUBMED	11042159					
REFERENCE	3					
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Saeki, N., Carninci, P., Suno, H., Akiyama, J., Nishi, K., Kitasuna, T., Tashiro, H., Itoh, M., Suni, L., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Masumoto, H., Sakaguchi, S., Ikegami, T., Kasaiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.					
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer					
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)					
MEDLINE	20530913					
PUBMED	11076861					
REFERENCE	4					
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.					
TITLE	Functional annotation of a full-length mouse cDNA collection					
JOURNAL	Nature 409, 685-690 (2001)					
REFERENCE	5					
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.					
TITLE	Analysis of mouse transcriptome based on functional annotation of 60,770 full-length cDNAs					
JOURNAL	Nature 420, 563-573 (2002)					
REFERENCE	6					
AUTHORS	(bases 1 to 4927)					
TITLE	Adachi, Y., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hizomoto, K., Hisatsaka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Higawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kohjiya, Y., Kondo, S., Komori, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, K., Sekai, C., Sekai, K., Sekazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Takabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takeku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.					
REFERENCE	Muramatsu, M. and Hayashizaki, Y.					
AUTHORS	Direct Submission					

QY 1729 GGACGACGACATCCCTGATCTG 1754
 |||||
 Db 1644 GGACCAAGACTCTTCCCGAGAGATG 1669

RESULT 9
 AK078953
 LOCUS 1958 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130231C15 product:similar to LIVER CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus], full insert sequence.

ACCESSION
 VERSION AK078953.1 GI:26347654
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Mech. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636

REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159

REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasahagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861

TITLE The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
 JOURNAL Functional annotation of a full-length mouse cDNA collection
 REFERENCE Nature 409, 685-690 (2001)
 AUTHORS 5

TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 JOURNAL Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 REFERENCE Nature 420, 563-573 (2002)
 AUTHORS 6 (bases 1 to 1958)

TITLE Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashizaki, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnate, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of

COMMENT
 Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.
 Please visit our web site for further details.
 URL: <http://genome.gsc.riken.jp/>.
 URL: <http://fantom.gsc.riken.jp/>.
 Location/Qualifiers
 1..1958

FEATURES
 source
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="PANTOM.DB:9130231C15"
 /db_xref="taxon:10090"
 /clone="9130231C15"
 /sex="male"
 /tissue_type="cecum"
 /clone_id="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 66..1742
 /note="unnamed protein product; putative similar to LIVER CARBOXYESTERASE PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus] (SWISSPROT|Q64419, evidence: FASTY, 78..418D, 100%length, match=1674)"
 /codon_start=1
 /protein_id="BAC37476.1"
 /db_xref="GI:26347655"
 /translation="MPLARLPGMLCVAGCLLLIOLHYGSDASPIRINTRGVRS FVNVKDTKSGVAFGLIPAPKPPVGLRPAPEDEPSPSGVDRDTSOPAMCLOPIYN LBAKAMNLIPLPISSEDCYLNTYTPHROESNLVPMWTHGGGLVGSASANDV SKLAETSEIIVIAIOLRGVGFPGSTGOHARGNGYIDVPAALRWVKNTAIFGGRN DRYTIFGVASGTGVSYSHTLSFMSKGLTHGAIMOSGVALLDLSISTSEVYKTVAN SCGEANTDESALHCHIRAKSKOSTLAINOVFNKI PAVVDGEFLPKRPOBLTSMFHVAN PSLIGVNTDECGWGMFMGDLHI IKNTRETLPAVLKTAAPMLPPECGLLVEEY WEDTDPEPTLOQVFEMLGDMEFVLPALQVAFORSQAPVAFYEFORHLSFTKIVRS HYKAAHGDAVAFVGSYLYMDNMLDTEEEELKRMKMYMANPARNGSPNEGLSPNP VLDHBOYLQDLPDPVAVGRALKARLQFWTTLTPQKIQELKSGSDKRAEL"
 1937..1942
 /note="putative"
 1958
 /note="putative"
 polyA_site
 polyA_site
 /note="putative"

ORIGIN
 Query Match 17.8%; Score 383.6; DB 3; Length 1958;
 Best Local Similarity 55.1%; Pred. No. 1.3e-92;
 Matches 855; Conservative 0; Mismatches 664; Indels 33; Gaps 4;

QY 210 AGGAACACGAGCTGGGATGATTCAGGCGAAGCAAGTCACTGTGCGAAGCCCTGTG 269
 |||||
 Db 165 AGGAACACCAAGGCGTCAGAGTGAAGGCGAGCTTGTGCCAGTGAAGCACTAAATCT 224

QY 270 CCTGTGAACGTGTTCTCGAGAGTCCCTTTGCTGCTCCCGCGTGGATCCCTGGATTT 329
 |||||
 Db 225 GGTGTCAATGCTTCTCGGAATATCCCTTTGCAAGCTCTCTTGAAGCACTAGCGGCTTT 284

QY 330 ACGAACCGGACGCTGCATGCGCTCGGATTAATCTTGAGAGAAGCCACTCTCACTCAAT 389
 |||||
 Db 285 GCACTCTCTGAGACCTGTGAGCCATGAGTGGTGTGAGATGAGACCTCAAGCCGGCC 344

QY 390 TTGTGCTCCGAGACTGAGATGCTGCTCTTGAATCAACATGCTCAAGTGCAATTAC 449
 |||||
 Db 345 ATGTGTCTGACGCTGATTAATTAATTTAGAGAAATCGGAAGAAATGATCGATCTG 404

QY 450 CCGAAATTTGGAAGTGCAGAGACTGCTCACTGAACATCTAATGCGGCGCCAGGCC 509
 |||||
 Db 405 CTTCCACTCTCTAATGTCCGAGGACTGCTGTATCTCAACATCTACACCAACCCAGGCC 464

QY 510 GATACAGCTCCAGCTCCCGCTTGGTGTGTTCCAGAGAGTGCCTTCAAGACTGGC 569
 DB 465 CAGAGAGGCTCTAACTGCTGCTGATGTGATGATTCACAGGTGTGACTGGTTGATAGA 524
 QY 570 TCAAGCTTCATTTGATGGGTCCGCTGGCTCTATGAGAGACGTGTGTGTGTC 629
 DB 525 TCAGCTTCATGAAATGATGATCTAACTGGCAGCCACTGAGGAATAGTAATTTGTCGC 584
 QY 630 GTCCAGTACCGGCTGGAATATTTGGTTCTTCAACAATGGGATCAGATGCTCCGGG 689
 DB 585 ATCCAGTATCGCTGGGGTCTTGGCTTTCAGACACTGAGACAGATCCAGAGGC 644
 QY 690 AACTGGGCTTCAAGAGCAAGGTGGCTGCTGTCTGGGTCCAGAAAGACATCGAGTTC 749
 DB 645 AACTGGGATACCTGAGACCAAGTGGCTGCTTAAGCTGGGTGACAAAGATTTGCTTAC 704
 QY 750 TTCCGTGGGAGACCCCAAGCTCTGTGACCATCTTGGCGAGTCCGCGAGCCATAAGTGT 809
 DB 705 TTTGAGGCAACCGTACAGGGGTCACTAATTTGAGTCTCAGCAGAGGTCAAGTGTG 764
 QY 810 TCTAGCTTTAATCTGTCTCCATGCGCAAGGCTTATTCACAAGCCATCATGAGAGT 869
 DB 765 TCTTCTAATATCTGTCTCCCAATGTCCAAAGGACTTTTCATGAGCCATTATGACAGT 824
 QY 870 GGGGTGGCATCATCCCTTACCTGAGAGGCCATGATTAAGAGAGTGAAGACCTGCAG 929
 DB 825 GGAATGGCCCTGCTGCTGACTCTCATCTGTACACTTCTGAG-----GTGCTTACAG 878
 QY 930 GTGGTTGACATTTCTGTGTAAACATAGCGTCAACTGTGAGCCCTGTGAGTGCCTG 989
 DB 879 ACAGTAGCAACCTATCTGTGTGAGAGCAGACACTCAAGAGCCCTGATTCAGTGCCTG 938
 QY 990 AGAGCAAAACCTCCAGAGAGTGTGACCTCAGCCAGCAAAACAAAGCTTTCACCTGA 1049
 DB 939 AGAGCAAAAGATMAACAAGAGATCTGTCTAATTAACAGGTCTTCAAGATGATCCCTG 998
 QY 1050 GTGGTTAGTGTCTTCTTCTTCTAATGAGCTCTAGATCTATGTCTCAGAAACATTT 1109
 DB 999 GTGGTGAATGGGAGTCTCTACCAAGACATCTCAGAGAGCTCTACCTCTATGATTTT 1058
 QY 1110 AAAGCAATCTCTTCATCATGAGATCAATAACAAGATGTGCTCTGCTGCTATG 1169
 DB 1059 CACCTGTCCCCAGCATATGTTGTGTCAACAAGATGATGTGCTGGGAGATCCCATG 1118
 QY 1170 AAGAGGCTCTG-----AGATCTCAAGTGTCTCAACAAGTCCCTTGCCTC 1217
 DB 1119 TTCAATGGGCTTGAACCATATCATMAAAGAACATMAACAGAGAGACCTGCTGCTTCTG 1178
 QY 1218 CATCTGAATACAAACATCTGACATCCCGCTCAGTATTTGACACTGTGTGCTAATGA 1277
 DB 1179 AAGAACAGAGCAGCAGATGATGTTGCTCCAGATGATGATCACTGCTGATGAGAG 1238
 QY 1278 TACTTCCATGACAAAGCACTCCCTGACTGAATCCAGACAGCTTCTTGGACTTGTGGA 1337
 DB 1239 TATATGGGGGACACTGAGAGACCCAGAACCTCTGCAAGACAGTTGAGAGATGTGGGG 1298
 QY 1338 GATGTGTTCTTTGTGTGCTCTGCACTGATCAGAGTCAATATCAGAGATGCTGTGCA 1397
 DB 1299 GACTTCATGTTGTGATCCCTGCTC---TCCAAATGGGACATTTTCAGCGTCCAGAGCT 1355
 QY 1398 CCTGTCTACTCTAATGATTTTCGACCGGCTCAGTCTTGTGAAGACAGAGCGGCT 1457
 DB 1356 CCTGTCTACTCTAATGATTTTCAGCATCTGTCAAGCTTCAACAGATGTCAGGCAATCA 1415
 QY 1458 TTTGTCAAGCGGACCAAGCTGATGAAGTCCGCTTGTGTGTGCTGTGCTTCTCTGAAG 1517
 DB 1416 CATGGAAGGAGACACATGTGATGACGTGTGCTTGTGCTTGTGCTCATTTTGGGGAGC 1475
 QY 1518 GGGGACATTTGATTTGTAAGAGACCAAGAGAGAGAGAGATTAAGTGAACCGGAGATG 1577
 DB 1476 ATGAATTTGAGCT-----CACTGAGGAGGAGAGCTCTAAGAGATGATG 1523

QY 1578 ATGAATATCTGGGCTACCTTGTCTGAAACCGGAATCTTAATGGAGCAGCTGTCTG 1637
 DB 1524 ATGAATATCTGGGCTACCTTGTCTGAAATGGAGAACCCCAACAGATGAGGGTCTACCTCC 1583
 QY 1638 TGGCCAGCTTATATCTGACATGACATCTCCAGCTGTGACTTGAACATGAGCTTCGA 1697
 DB 1584 TGGCTGTGTGGACCATGATGAGCAGTACCTGACGCTGGAACCCAGCTGTGTGGGT 1643
 QY 1698 CAGAGACTCAAGAACCGCGGTGATTTTGGACCAAGACCATCCCTGA 1749
 DB 1644 CAGGCTTGAAGGCCAGAAAGCTGCAATTTGACCAAGACTCTGCCCAAG 1695

RESULT 10
 AK078879
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AK078879 1944 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030604P03 product:carboxylesterase 3, full insert sequence.
 AK078879
 AK078879.1 GI:26347580
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLES
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 3
 4
 5

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 4
 5

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsuna, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, N., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 5

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 6

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1944)
 Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

TITLE
JOURNAL

Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akhiba, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

COMMENT

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES
Source

1. 1944
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:9030604P03"
/db_xref="taxon:10090"
/clone="9030604P03"
/sex="male"
/clone_type="colony"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
64..1761
/note="unlabeled protein product; carboxylesterase 3 (MGD|MG1:2148202, GB|NM_053200, evidence: BLASTN, 99%, match=1939)
putative"
/codon_start=1
/protein_id="BAC37439.1"
/db_xref="GI:26347581"
/translation="MGILPLIWLISLACTAMGYPSPVNVTVKGVLGKYNLEGPQVAVFLGVPFAKPLPLSLRFAFPQAPAPSPVNTKTSYPMCSQDAVGSQVLSLPTNKENIPLQPSDECLYNIYTPADLTAKSRPLPVMVWHIGGLVVGASSTYGLSLSHENVVVTIOYRLGIVGFPSTDSHRSGWGHLDVAAIRWQDNIANFGAPSVTLFSSAGAGFSVSLVYIWSPIAKNLPKRAISSGVSLTAALITTVKRIAGLVATLSCGKTTLSAIVWHICRQTEDELLETSLKLNLFLLDLGNPKESVPLPTVIGVVLGKCKTTLAESFSTVPTIVGINKQEPFWIIPITLMGVPLARGLDQKTNANLIMKSYPTLKISENMFVVAEKYLGSTDLTLTKKDLFDPLMDVVGVPVSIVSRSHDAGASTYMEFEYRPSFVAMRPKAVIGDHGDEIFSVFSGPLKDGASEEETNLSKVMKFMANPARGNNGGGLPHMPEYDOKEGYLKIAGSTQAQRLKDXEVSFMAELRAKESAPRSHREVEL"
1922..1927
/note="putative"
1944
/note="putative"

ORIGIN

polyA_signal
polyA_site
/note="putative"
1944
/note="putative"

Query Match 16.5%; Score 355.2; DB 3; Length 1944;
Best Local Similarity 55.1%; Pred. No. 7e-85;

Matches 870; Conservative 0; Mismatches 633; Indels 75; Gaps 6;

QY 212 GAACACCAAGGTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAAAGCCCTGGCC 271
DB 141 GAACACCTTTAAAGCAAGTCTGGGGAGTAGTCAATTGGAGGATTCAACAGCC 200
QY 272 TGTGAAGCTGTTCTCTGGAAGTCCCTTGTGCTGCCCGCCGCTGGATTCCTGCGATTAC 331
DB 201 TGTGGCTGTTTCTCTGGGAAGTCCCTTTGCCAAGCCCTCTTGTGCTCTTGGAGATTGC 260
QY 332 GAACCCGAGCTGCGATCGCCCTGGAGTAACTTGCAGAGAGCCACCTCTTACCTTAATT 391
DB 261 TCACCAACACCTGGACAGCCCTGGAGCTTGTGAAGAAACACACCTCTTACCCGCTTAT 320
QY 392 GTGCTCCAGAGCTC-----AGATGGCTGCTCTTATGATCAACACATGCTCAAGGTGA 445

DB 321 GTGCTCTCAGAGTACGCTTTGGAGGCGAGTGTCTTCAAGCTCTTACCAACAGAGAGA 380
QY 446 TTACCCGAATTCGGAGTGTCAAGAGTGTCTTACCTTACCTTGAACATCTATGCGCTGCCA 505
DB 381 GAACATTCCTTTACAGTTTTCGAAAGCTGCTTACCTGAATATTACATCTCTGCTGA 440
QY 506 CCCCCATACAGGCTCCAAAGCTCCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 565
DB 441 CTTCAGAAAGAACAGAC 500
QY 566 TGGCTCAGCTTCATCTTTGATGAGTGTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCT 625
DB 501 GGGCGAGACATCCACTTATGATGAGCTGCGCCCTCTGCTGCCATGAAAATGTGTGTGTGT 560
QY 626 GGTGTTCAGTACCGGCTAGGAATATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 685
DB 561 GACCTTCAGTATGCGCTTGGCATCTGGGATTCCTTACGACCTGGGGATGACACAGTGTG 620
QY 686 GGGGAAGTGGGCTTTCAGAGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 745
DB 621 GGGAAAGTGGGGTCACTTGGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 680
QY 746 GTTCTGTGGTGGGAGCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 805
DB 681 CAATCTTGGGGGCAACCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 740
QY 806 TGTCTTACGTCTTATCTGTCTGTCCCAAGGCTTATCCAAAGCCATCATGTGA 865
DB 741 TGTCTGTCTGT 800
QY 866 GAGTGGGTGGCATCATCTCTTACCTGAGGCCATGATTTAGAGAGTGAAGAGTGAAGT 925
DB 801 GAGTGGGTGTCTCTCACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 860
QY 926 GCAAGT 985
DB 861 GGT 920
QY 986 CTTGAGAGCA-----AAACCTCCAAAGAGGTG 1013
DB 921 CCAAGAGACAGAGATGAATCTAGAGACCTCACTAAATTTGAAATCTTTTAACTGGA 980
QY 1014 CTGACCTCAGGCAAGAAACAAG-----TCTTCACTGAGTGTGTGTGTGTGTGTGT 1066
DB 981 CTTACTGTGAATTCCAAAGAGAGCTATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1040
QY 1067 CTTTCTTAATGAGCTCTAGATCTATGTGTCTCAGAAACATTTTAAAGCAATTCCTTCAT 1126
DB 1041 TCTGCCAAAGGCAACAGAGAGATCTGTGTGAGAGAGTTTCAGCATGTCTCTCTCAT 1100
QY 1127 CATTGAGTCAATTAACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1177
DB 1101 AGTGGCATTAACAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1160
QY 1178 TCTTGAATCTCTAGT 1234
DB 1161 GCTGAAGCAAACTGAGCAAGAGAGAGCAATTCCTCTTGTGGAAGTCTTACCCAAC 1220
QY 1235 CCTGCAATTCCTGCTGATATTGTGCACTTGTGCTTATGAAATCTTCCATGACAAGCA 1294
DB 1221 ACTTAAATCTCTGAGATATGATTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280
QY 1295 TCTCCTGACTGAATTCGAGACAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1354
DB 1281 TGACCTACCAAAAAGAAAGACCTGTTCAGAGACTGTATGTGTGTGTGTGTGTGTGTGT 1340
QY 1355 CCTGTGACTGATCAAGCTGTGATATCAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGT 1414
DB 1341 CCATCATGATGT 1400
QY 1415 GTTGTGGCAACGGGCTCAAGT 1474

Db 1401 GTTTAGATGTCGCCAGCTTTGTATCGGCATGAGACCAAGCAAGTAATGAGACCA 1460
QY 1475 CGCTGATGAAGTCCGCTTTGTGTGGTGGTGCCTTCTCGAAGGGGACATTGTATGTT 1534
Db 1461 TGGTATGATGATCTTCTCACTATTTGGATCTCCATTTTAA----- 1502
QY 1535 CGAAGACCCACGAGAGAGAGAGATTAAGTGAAGGAGATGATGAAATACCTGGGCTAC 1594
Db 1503 AGATGCTGCTCAAGAAAGAGAGACCACTGACGAAGATGTGATGAAATCTGGGCGCA 1562
QY 1555 CTTTCTCGAAGCCGGGAATCTTAATGGAGACCTGTCTGTGGCCAGCTTAATATCT 1654
Db 1563 CTTTCTCGGAGATGGAAACCCCAATGTGAGAGGCTGCCCCACCTGGCAATATGACCA 1622
QY 1655 GACTGAGACGATCCTCGACCTGATGAAACATGAGCTCGGACAGAGACTCAAGAAC 1714
Db 1633 GAGGAGGGGTATCTGAAAGATTTGTGCTCTCACTAGGAGCCGAGGCTGAGAGCA 1682
QY 1715 GCGGGTGAATTTTGGAC 1732
Db 1683 AGAGTGAATTTTGGGC 1700

RESULT 11
HSM803013 2888 bp mRNA linear HTC 22-SEP-2004
LOCUS HSM803013
DEFINITION Homo sapiens mRNA; CDNA DKFZp434N0935 (from clone DKFZp434N0935).
ACCESSION AL713761
VERSION AL713761.1 GI:19584506
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2888)
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Weil, B., Amid, C., Oeinger, A., Fobo, G., Han, M. and
Wiemann, S.
COMMENT The German cDNA Consortium
CONSRM Direct Submission
TITLE Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
JOURNAL Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp434N0935) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneId=DKFZp434N0935
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/
location/Qualifiers

FEATURES
source 1..2888
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp434N0935"
 /db_xref="taxon:9606"
 /clone="DKFZp434N0935"
 /tissue_type="testis"
 /clone_lib="434 (anonymous: hhes3). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 /notes="cardoxylesterase 2 isoform 2"
 1..2888
 /gene="DKFZp434N0935"
 954..2777
 /gene="DKFZp434N0935"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAD28531.1"
 /db_xref="GI:19584507"
 /db_xref="UniProt/Swiss-Prot:O00748"

/translation="MTAQSRLPTPTPEPSQRTPLPQVQTERLGAALHICWTPDG
OPLGQORVRRQRTESPTWRILHRLARLAVACGLILVRCGGOSASPIRTHT
GOVGSILVHVGAAVOTFLGIPAKPPLGRLAPAPPEPSMGVGGDGTMLPTL
ODLTAVERSEPLISQPMPTSDMSRDCILSYTPHBSNPLPVMWIRGALVPG
MASLYDSMLALENVVVITITGLVGLFSPSTGRKHA TKNGVYLDQVAAALRVQNI
AHFGNPDVITIPESAGTSVSLVSPISQGLPHGAIMESGVALPELILASADVI
STVANILASCOVQSEALVGLRGSKSEILAIINPKMTIPGVVGLVLRPRLSILA
SADFPVPSIVGVNNNEFMGLPKVMRIYDIOKEMDREASQALQKMLTLPPTGL
DLREBYIGNDGPOTLQAOFOEMWADSMFYI PALQVHAFQCSRAPIVYEFROHPSW
LKNIRPMKADHYVETFEESOLSRQMKVYANPNPNNEGILPHPLPDQESQYL
QNLQIPAVGRALKAHRLQPMKALPQKIQELPEERHTEL"

ORIGIN
Query Match 16.0%; Score 345.8; DB 3; Length 2888;
Best Local Similarity 54.8%; Pred. No. 2,9e-82;
Matches 850; Conservative 0; Mismatches 617; Indels 84; Gaps 5;

211 GGACACCAAGCTGGAGATGATTCAGGCGAACAGATCACTGTCTGGGAACCTGTGTC 270
Db 1246 GGACACACACACAGGGGAGAGGTGCTGGGAGACTTGTCCATGTGAAGGCGCAATGCGC 1305
QY 271 CTGTGAAGCTGTTCTCGGAGTCCCTTTGCTGCTCCCCGGTGGATCCCTGCATTTA 330
Db 1306 GGGTCAAACTTCTCGGAATTCATTGTCGAAGCCACTCTAGGTGCTGCGCATTTG 1365
QY 331 CGAACCCGACGCTGCATCGCCCTGGAGTAACTTCGAGAACCACTCTCACTAATTT 390
Db 1366 CACCCCTGAGGCCCTCTGAATTTTGAATGTGTGATGAGGATGAAACCACTCCGCCA 1425
QY 391 TGTGCTCGAAGACTCAAGATGGCTGCTTTAGA--TCACACATGCTCAAGTGCAAT 447
Db 1426 TGTGTCTACAGAGACTCAACGAGTGAGTCAAGATTTCTTAGCGATTCACATGACCT 1485
QY 448 ACCGCAATTCGAGATGTCAAGAGCTGCTTACTGAACTTAATGCGCTGCCACG 507
Db 1486 TCCCTTCGACCTCAATGTCTGAGACTGCTGTACTCAGCATCAACGCCGCCCA 1545
QY 508 CGATACAGGCTCCAAAGCTCCCGCTTGTGTGTTCCAGAGAGTGCCTTCAAGACTG 567
Db 1546 GCCATGAAGGCTTAACTGCCGATGTGTGATGATCAAGTGTGTGCTGTTTGTG 1605
QY 568 GCTACGCTCCATCTTGAATGAGTCCGCCCTGCTGCTTATGAGACGCTGTGTTGTG 627
Db 1606 GGATGGCTTCCTGTATGATGATGATTCATGCTGGCTGCTTGAAGAACGTGTGTGCTCA 1665
QY 628 TGTTCATGATCCGGCTAGAGATTTGTTTCTTACCAATGAGATCAGATGCTCGG 687
Db 1666 TCATCAGTACCGCTGGGTGTCTCGGCTTCTTACGACTGAGACAGGACGCAACG 1725
QY 688 GGAACCTGGGCTTCAAGAGACAGAGTGTGCTGTCCGCGGTCCAGAAAGATCAGAT 747
Db 1726 GGAATCGGAGCTACTGAGACCAAGTGGCTGCACTAGCGTGGTCCAGAAATATGCGCC 1785
QY 748 TCTTCGTGGGAGCCCAAGCTGTGACCATCTTTTGGAGAGTCCGCGGAGCCATAGTG 807
Db 1786 ACTTTGAGGAGCAACCTGACGCTGACCATTTTGGGAGTCTGCGGTGCGAGAGTG 1845
QY 808 TTTCTAGTCTTAATCTGTCTCCATGCGCAAGGCTTAATTCACAAAGCATCATGAGAG 867
Db 1846 TGTCTTCCTGTTGTGTGTCCTCCCATATCCCAAGGACTCTTCCACGAGCATCATGAGAG 1905
QY 868 GTGGGTGGCCATCATCTCCCTTAACCTGAGGCGCCATGATTAAGAGAGTGGAGACTGC 927
Db 1906 GTGGGTGGCCCTCTGCGCGGCTCATATGCGAGCTCAGCTGA-----TGTCACTTCCA 1959
QY 928 AGGTGTTGACATTTCTGTGTAAACAATGCGTCAAGCTTGAAGCCCTGTGAGTGTGC 987
Db 1960 CGGTGTGCCCAACTGTCTGCTGTGACCAAGTGAATCTGAGGCGCTGTGGCTGCC 2019
QY 988 TGAAGCAAAACCTTCAAGAGAGTGTGACCTTCAAGCAAGAAACAAGTCTTCACTC 1047
Db 2020 TCGCGGGGCAAGATGAAGAGATTTTGAATTAACAAGCTTCAAGATGATGCCG 2079

Db 1714 TCATCCAGTACCGCTGGGTGTCTTGCAGCATGAGCAAGCAAGCAAGCG 1773
 Qy 688 GGAATGGGCTTCAAGGACGAGTGGCTCTTCCGGGTCCAGAAACATCGAGT 747
 Db 1774 GCAATGGGGCTCACTGGACCAAGTGGCTGCACTAGCGGTCCAGCAAGATATGCCCC 1833
 Qy 748 TCTTCGATGGGGAACCCGAGCTCTGACATCTTGGGAGAGCCGCGGAGCCATAGTG 807
 Db 1834 ACTTTGAGGCAACCTGACCGGTGACCATTTTGGGAGGTCTGGGGTGGCAAGATG 1893
 Qy 808 TTTCTAGTCTTATCTGTCTTCCATGCGCAAGGCTTATTCACAAGCCATGAGAGA 867
 Db 1894 TGTCTCCCTTGTGTGTCTCCCATATCCCAAGGACTCTTCCACGAGACCATCATGAGA 1953
 Qy 868 GTGGGGTGGCCATCATCCCTTACCTGAGGCGCCATGATTAGAGAAGTGAAGCTGCG 927
 Db 1954 GTGGGGTGGCCCTCTGCGCGCTCATATGCGAGCTCAGCTGA-----TGTCAATCTCA 2007
 Qy 928 AGTGTGTGCAATTTCTGTGGTAACAATGCGTCAGACTGAGGCGCTGAGAGTCC 987
 Db 2008 CGGTGTGGCCCACTGTCTGCTGTGACCAAGTTGACTCTGAGGCGCTGCGGTGCG 2067
 Qy 988 TGAGGACAAACCTCTCAAGAGCTGTGACCTCAGCCAGAAACAAAGTCTTCACTC 1047
 Db 2068 TGCGGGGCAAGATTAAGAGGATTTCTTGCAATTAACAGGCTTCAAGATGATCCCG 2127
 Qy 1048 GAGGTGTGATGAGTCTTCTTCTTCTTATGAGCCTTATGATCTATTTGTCTAGAAAGAT 1107
 Db 2128 GAGGTGTGATGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2187
 Qy 1108 TTAAGCAATTTCTTTCATCATCGGAGTCAATTAACACAGAGTGTGGTCTTCTGCTGCC-- 1165
 Db 2188 TTTGAGCCCTGTCCCTAGCATGTGTGTGTGCAACAAATGATGCGCTGCTCATCCCA 2247
 Qy 1166 -----TATGAGAGAGGCTCTGAGATCTTCAAGTGTCTCAACAAATCTTGTGCC 1215
 Db 2248 AGGTATGAGAGATCTATGATACCCAGAAAGAAATGAGACAGAGGCTCTCCAGGCTGCTC 2307
 Qy 1216 TCCATCTGATACAAACATCTCTGACATCCCGCTCAGTATTTGACCTTGTGGTATATG 1275
 Db 2308 TGCAGAAATTTTAAAGCTGTGATGTGTCTTCTTCAATCTGTGTGACCTGTGAGGAGG 2367
 Qy 1276 AATATCTCATGACCAAGCACTCCCTGACTGAAATCCGAGACAGTCTTGTGACTTGTG 1335
 Db 2368 AGTACATTTGGGACAAATGGGATCCCAAGACCTTCAAGCGAGTTCAGAGAGATGATGG 2427
 Qy 1336 GAGATGTGTCTTGTGTGTCTCTGCACTGATCAAGCTGATATCAAGAGATGCTGTG 1395
 Db 2428 CGGACTCATGTGTGTGTATCCCTGCACTCAAGTAGCATTTTCAAGT---TTCGCGGG 2484
 Qy 1396 CACCTGTCTACTTCTATGAGTTTGGGACCGGCTCAGTGTCTTGAAGAACAAGCCGG 1455
 Db 2485 CCCCTGTGTACTTCTACAGATTTCCAGCATCAGCCAGCTGTGCTCAAGAAACATCAGGCCAC 2544
 Qy 1456 CTTTGTCAAGCCGACCAAGCTGATGAAGTCCGTTTGTGTGTGTGTGTGTGTGTGTGTGT 1515
 Db 2545 CGCAGATBAAGCCAGCATTTAA----- 2569
 Qy 1516 AGGGGGAATTTATTTATTCAGAGAGCCACGAGAGAGAGAGATTAAGTCAAGCCGAGAA 1575
 Db 2570 -----ATTCACTGAG 2604
 Qy 1576 TGAATGAATACTGGGCTACTTTTGTCTGAAACCGGGAATCTTAATGGAAACGACTGTCTC 1635
 Db 2605 TGAATGAATACTGGGCTCAACTTTTGGAGAAATGGGAACCCCAATGCGAGAGGTCTGCAAC 2664
 Qy 1636 TGTGGGCGAGCTTATTAATCTGACTGAGAGTACTCTCAGCTGAGACTTGAACAATGAGCTCG 1695
 Db 2665 ACTGGCGCTGTGCAACAGAGAGAGCAATACTCGGCTGAGCTTAAGCTGTGGGCTGG 2724
 Qy 1696 GACAGAGACTCAAGAAACCGCGGGTGAATTTTGGACCAAGCAATCCGCC 1746
 Db 2725 GCGGGGCTCTGAAGGCCCAAGGCTCCAGTTCTTGAAGAGAGCGGTGCCCC 2775

RESULT 13
 CFI11083
 LOCUS
 DEFINITION
 Shultzomic04334 Rat lung airway and parenchyma cDNA libraries
 Rattus norvegicus cDNA clone Contig3845 5', mRNA sequence.
 CFI11083
 ACCESSION
 Rattus norvegicus
 KEYWORDS
 EST.
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 1918)
 Shultz,M.A., Zhang,W., Gu,Y.-Z., Baker,G.L., Fannuchi,M.V.,
 Padua,A.M., Gurske,W.A., Morin,D., Penn,S.G., Jovanovich,S.B.,
 Plopper,C.G. and Buckpitt,A.R.
 Gene expression analysis in response to lung toxicants: I.
 Sequencing and microarray development
 Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)
 CONTACT
 Contact: Shultz, MA
 Dept. of Molecular Biosciences, School of Veterinary Medicine
 University of California, Davis
 1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA
 Tel: 530 752 0793
 Fax: 530 752 4698
 Email: mshultz@ucdavis.edu
 Average Phred score is 20 or better. All poor quality data (Phred <
 20) and vector/linker sequence has been removed.
 High quality sequence atop: 1918.
 FEATURES
 source
 1..1918
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="Contig3845"
 /sex="male"
 /tissue_type="airway or parenchyma"
 /dev_stage="adult"
 /clone_lib="Rat Lung airway and parenchyma cDNA libraries"
 /note="Organ: Lung; Vector: pGEM-11zf(-); Site 1: Eco RI;
 Site 2: Not I; mRNA was isolated from microdissected rat
 lung airways and parenchyma tissues."

ORIGIN
 Query Match 15.7%; Score 338; DB 7; Length 1918;
 Best Local Similarity 54.2%; Pred. No. 3.4e-80;
 Matches 869; Conservative 0; Mismatches 660; Indels 75; Gaps 6;
 Qy 186 GGGCTTGTGTCGTAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGCAAGCA 245
 Db 83 GGGATATCATCTCTACCACTGTGTGTGAACACTGTAAAGCAAGTCTGGGAAATAT 142
 Qy 246 GTCACTGTGCTGGAGAGCCCTGTGTGTGAACGTGTCTCTGAGAGTCCCTTGTCTGT 305
 Db 143 GTCAATTTGGAAGATTTGCAAGCTGTGTGTCTTCTTCTGGGAAATCCCTTCCCAAG 202
 Qy 306 CCCCCGTGGATCCCTGCGAATTAAGAACCGGACCTGATGAGCCCTGGGATTAATTTG 365
 Db 203 TCCCTCTTGGGCTCTTGAAGTTGCTCTCAACACAGCTGTGAAGTCTTGAATTTGTG 262
 Qy 366 CGAAGGCACTCTCTACCTTAATTTGTGCTCCAGAACTC-----AGAGTGTGCTC 419
 Db 263 AAGAATTAATCTCTTACCTACCATATGCTCTCAAGATGCTGTGAGGCAAGGTTCTC 322
 Qy 420 TTAATCAACAGCTTAAGTGCATTAACCCGAATTCGAGAGTCAAGAGCTCAAGCTCTC 479
 Db 323 TCAGAGCTTTTCAACCAAGAGAGAAACATTCCTTTACAGTTTCTGAAGACGTGCTC 382
 Qy 480 TACTGAACATCTATGCGCTGCCAGCGATACAGGCTCCAAGCTCCCGTCTTGTGTG 539

Db	383	TACCTGAATGTTATATCTCCCGCTGACTTGACAAAGAACGCCGGTTACAGATGATG	442	Db	1463	TTTTTAA-----AGATGCCCTTCAGAAAGAGACCAATCTCAGC	1504
Qy	540	TGGTTCACAGAGAGTGCCTTCAAGACTGGCTACGCTCCATCTTGAAGGTCGCCCTG	599	Qy	1569	CGGAAGATGATGAATATCTGGGCTACCTTTCCTGAAACCGGAAATCTTAATGGAAAGAC	1628
Db	443	TGATATCAATGAGAGTGAAGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	502	Db	1505	AAATGTGATGAATATCTGGGCAACTTTGCTCGAATGGAAACCTTAATGGGGAGGG	1564
Qy	600	GCTGCTATGAGGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	659	Qy	1629	CTGTCTGT	1688
Db	503	TCTGCCCATGAATAATGT	562	Db	1565	CTGCCCATTTGGCCAGAAATATGACCAAGAGAGAGGATCTCGAAGATGTGTGCTCAACT	1624
Qy	660	TTCAACCAATGGGATCAGATGCTCCGGGAACTGGGCTTCAAGGACCAAGGTGTGTGT	719	Qy	1689	AGCCTCGACAGAGACTCAAGAAACCGCGGTGATTTTGGAC	1732
Db	563	TTCAAGACAGGGGATGAAACAGCCAGGGCACTGGGGTCAATTGGACAGCTGGGTGCA	622	Db	1625	CAGGAGGCCAGAGGCTGAAGAGCAAGAAAGTGGCTTTTGTC	1668
Qy	720	CTGTCTGGGTCCAGAAACATCCAGTCTTCGGTGGGGAACCCAGCTCTGTGACATC	779	RESULT 14			
Db	623	CTGCACGTGGGTCCAGGACAAATATGCAACTTTGGGGGCAACCCAGGCTCTGTGACATC	682	LOCUS	AK007235		
Qy	780	TTTGGCGAGTCCGCGGAGCCATAAGTGTCTTCTAGCTTATATCTGTCTCCATGGCAAA	839	DEFINITION	AK007235	989 bp mRNA linear HTC 03-Apr-2004	
Db	683	TTTGGAGATCTGCAAGGATTTCAAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	742	VERSION	AK007235	GI:12840656	
Qy	840	GGCTTATTTCAAAAGCCATCATGGAAGTGGGGTGGCCATCATCTTACCTGGAGGCC	899	KEYWORDS	AK007235.1	HTC; CAP trapper.	
Db	743	AACTCTTCCACAGGGCCATTTCTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	802	SOURCE	Mus musculus	(house mouse)	
Qy	900	CATGATTAATGAGAAAGTGAAGACCCGACGATGTGTGACATTTCTGTGTGTAAACAATCG	959	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Db	803	ACAGATACCAAGCCCATTTCTTAATGTGATGTCTACTCTTTGGGTGTAACCAACACACA	862	REFERENCE	Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
Qy	960	TCAGACTGTAGGCTCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	995	AUTHORS	Carninci, P. and Hayashizaki, Y.		
Db	863	TCAGCTGTATGTTCATTGCTGTGGCCAGAAAGACAGAGATGAATCTCGGAGACTTCA	922	TITLE	High-efficiency full-length cDNA cloning		
Qy	996	-----AAACCTCCAGAGAGTGTGTGACCTTCAGCCAGCAAAAGTCT-----	1040	JOURNAL	Meth. Enzymol.	303, 19-44 (1999)	
Db	923	TTAAATGAAATCTTTTAACTGAGACTTATTTGGAAACCCAAAGAGACTATCTCCCTTC	982	PUBMED	99279253		
Qy	1041	TTCACTCGAGT	1100	REFERENCE	10349636		
Db	983	CTACTACTGTGATGATGAGT	1042	JOURNAL	20499374		
Qy	1101	AAAGCATTTAAAGCAATCTTCCATCATCGAGTGAATTAACAAGAGTGTGCTTCTCTG	1160	PUBMED	11042159		
Db	1043	AAAGATTTCACAGGGTCCCTACATAGTGGGCAATCAACAAGAGATTGGCTGGATC	1102	REFERENCE	3		
Qy	1161	CTGCCTATG-----AAGAGGCTCCGAGATCTCTGATGTGTGTGTGTGTGTGTGTGT	1210	AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,		
Db	1103	ATTCATATGCTTAAAGGCTATCTCTCCGAAAGCAACTGGAACAGAAACAGCCAAAT	1162	TITLE	Kono, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M.,		
Qy	1211	TGCCCTCATGTGAACAAAC--ATCTGCAATCCGCTCTCATGTAATTTGCACTTGTG	1268	JOURNAL	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
Db	1163	TCCCTCTGTGAGAGTCTTACCAACTGAAATCTCTGAAATGATTTCCAATGGCT	1222	PUBMED	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
Qy	1269	GCTAATGAATCTTCCATGACAGACCTCCGTGACTGAATCCGAGACAGTCTTGTGAC	1328	REFERENCE	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,		
Db	1223	TCTGAAGATATTTAGAGGAGCAGATGACCTGCCAAGAAAGAACTGTTCAGAGAC	1282	AUTHORS	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,		
Qy	1329	TTGCTTGAAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1388	TITLE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
Db	1283	TTGGTTCAGATGTGATTTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1342	JOURNAL	RIKEN integrated sequence analysis (RISA) system-384-format		
Qy	1389	GCTGTGCACTGTCTTCTATGATTTGGGACCGGCTCTCATGCTTTGGAAGACG	1448	PUBMED	Genome Res.	10 (11), 1757-1771 (2000)	
Db	1343	GCTGGAGCCCCCACTTCAATGATGAATTTGATGCGCCCAAGCTTTGATACGACATG	1402	REFERENCE	20530913		
Qy	1449	AAGCGGCTTTTTCAGAAAGCCAGCGCATGATGAATGTCGCTTTGTGTGTGTGTGTGTGT	1508	AUTHORS	11076861		
Db	1403	AGGCCCAAGACGATCGAGACCAATGTGATGAATTTCTTCAATTTTGAATTCCTCA	1462	TITLE	4		
Qy	1509	TTCTGAAGGGGGAATTTTATGTTTGAAGAGCCAGGAGAGAGATTAAGTACGAGC	1568	JOURNAL	The RIKEN Genome Exploration Research Group Phase II Team and the		


```
OY 301 CTGCTCCCCGCGTGGAGATCCCTGCGATTTACGAACCCGCGAGCTGATCGCCCTGGAGATA 360
Db 301 CTGCTCCCCGCGTGGAGATCCCTGCGATTTACGAACCCGCGAGCTGATCGCCCTGGAGATA 360
OY 361 ACTTGGAGAGAGCCACCTCTACCCCTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTCT 420
Db 361 ACTTGGAGAGAGCCACCTCTACCCCTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTCT 420
OY 421 TAGATCAACATGCTCAGAGTGCATTAACCCGAAATTTGGAGTGTCAAGAACCTGCTCT 480
Db 421 TAGATCAACATGCTCAGAGTGCATTAACCCGAAATTTGGAGTGTCAAGAACCTGCTCT 480
OY 481 ACCTGAACATCTATGCGCTGCGCCGCGCATACAGGCTCCAGAGTCCCGCTTGTGTGT 540
Db 481 ACCTGAACATCTATGCGCTGCGCCGCGCATACAGGCTCCAGAGTCCCGCTTGTGTGT 540
OY 541 GGTTCGCGAGAGTGGCTTTCAAGACTGCTCAAGCTCCATCTTGTATGGGTCCGCTGG 600
Db 541 GGTTCGCGAGAGTGGCTTTCAAGACTGCTCAAGCTCCATCTTGTATGGGTCCGCTGG 600
OY 601 CTGCTATGAGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 601 CTGCTATGAGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
OY 661 TCACCACATGGGATCAGATGCTCCGCGGAACTGGGCTTCAAGAACAGGTGGCTGCTC 720
Db 661 TCACCACATGGGATCAGATGCTCCGCGGAACTGGGCTTCAAGAACAGGTGGCTGCTC 720
OY 721 TGTCTGTGGTTCAGAGAAATCAGATTTCTCGGTGGGAGCCCACTGCTGTGACATCT 780
Db 721 TGTCTGTGGTTCAGAGAAATCAGATTTCTCGGTGGGAGCCCACTGCTGTGACATCT 780
OY 781 TTGGCGAGTCCGCGGAGCCCAATAGTGTCTAGTCTTATAGTCTCCATGCGCAAG 840
Db 781 TTGGCGAGTCCGCGGAGCCCAATAGTGTCTAGTCTTATAGTCTCCATGCGCAAG 840
OY 841 GCTTATTCACAAAGCATCATGAGAGTGGGTGGCCATCTCTTACCTGAGGCC 900
Db 841 GCTTATTCACAAAGCATCATGAGAGTGGGTGGCCATCTCTTACCTGAGGCC 900
OY 901 ATGATTAATGAAAGATGAGACCTGCAAGTGTGCAATTTCTGTGTAACTATGCT 960
Db 901 ATGATTAATGAAAGATGAGACCTGCAAGTGTGCAATTTCTGTGTAACTATGCT 960
OY 961 CAGACTCTGAGGCGCTGCTGAGAGTGTGAGAGCAAAACCTCCAGAGAGCTGAGACC 1020
Db 961 CAGACTCTGAGGCGCTGCTGAGAGTGTGAGAGCAAAACCTCCAGAGAGCTGAGACC 1020
OY 1021 TAGCCAGAAACAAAGTCTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1021 TAGCCAGAAACAAAGTCTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
OY 1081 CTCTAGATCTATGTCTGAGAAAGATTTAAAGCAATCTTCCATCATCGAGTCAATA 1140
Db 1081 CTCTAGATCTATGTCTGAGAAAGATTTAAAGCAATCTTCCATCATCGAGTCAATA 1140
OY 1141 ACCAGAGTGTGCTTCTGCTGCTATGAGAGAGCTCCGAGATCTCAGTGGCTCCA 1200
Db 1141 ACCAGAGTGTGCTTCTGCTGCTATGAGAGAGCTCCGAGATCTCAGTGGCTCCA 1200
OY 1201 ACAAGTCCCTGCGCTCATCTGATACAAACATCTGCAATCTCCGCTCAAGTATTGG 1260
Db 1201 ACAAGTCCCTGCGCTCATCTGATACAAACATCTGCAATCTCCGCTCAAGTATTGG 1260
OY 1261 ACCTGTGTGCTAATGATTAATCTTCACTGACAGAGCTCCGAGTGAATCCGAGAGCTC 1320
Db 1261 ACCTGTGTGCTAATGATTAATCTTCACTGACAGAGCTCCGAGTGAATCCGAGAGCTC 1320
OY 1321 TTCTGAGCTTGTGTGAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
Db 1321 TTCTGAGCTTGTGTGAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
```

```
OY 1381 ACAGAGATGTGTGCACTGTCTACTTATGATGTTTCGAGCCGCTCAGTGTCTTG 1440
Db 1381 ACAGAGATGTGTGCACTGTCTACTTATGATGTTTCGAGCCGCTCAGTGTCTTG 1440
OY 1441 AAGACAGAGAGCGCTTGTCTCAAGCGGACAGCTGATGAAGTCCGCTTGTGTCTG 1500
Db 1441 AAGACAGAGAGCGCTTGTCTCAAGCGGACAGCTGATGAAGTCCGCTTGTGTCTG 1500
OY 1501 GTGTGCTTCTTCAAGAGGAGGACATTTGTATGTTGAAGAGCCAGAGAGAGAGAGT 1560
Db 1501 GTGTGCTTCTTCAAGAGGAGGACATTTGTATGTTGAAGAGCCAGAGAGAGAGAGT 1560
OY 1561 TACTGAGCGGAGATGATGAATTACTGGGCTACCTTTGCTCGAACCAGGAACTCTATG 1620
Db 1561 TACTGAGCGGAGATGATGAATTACTGGGCTACCTTTGCTCGAACCAGGAACTCTATG 1620
OY 1621 GGAAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
Db 1621 GGAAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
OY 1681 TGAACATGAGCTCGGACAGAGACTCAAGAAACCGCGGTGTGATTTTGGACAGACCA 1740
Db 1681 TGAACATGAGCTCGGACAGAGACTCAAGAAACCGCGGTGTGATTTTGGACAGACCA 1740
OY 1741 TCCCTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
Db 1741 TCCCTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
OY 1801 TCTCTCTCTCGAGCCCTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
Db 1801 TCTCTCTCTCGAGCCCTTCTTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
OY 1861 TTGTGTTCCTTCTCTCTCCATTAATTTCTCCGCAATATTAGCTTCTTTGTAGCTCA 1920
Db 1861 TTGTGTTCCTTCTCTCTCCATTAATTTCTCCGCAATATTAGCTTCTTTGTAGCTCA 1920
OY 1921 GCGCTTCTCATAGGAGATCCTTGCAAAACAAAGCTGCTTCCGTGAATTTATGACTTA 1980
Db 1921 GCGCTTCTCATAGGAGATCCTTGCAAAACAAAGCTGCTTCCGTGAATTTATGACTTA 1980
OY 1981 GGAATATCTTACAGAAATCTTTTCAACATCAAAAGTGCATTTGTCTTGGAGAGGAA 2040
Db 1981 GGAATATCTTACAGAAATCTTTTCAACATCAAAAGTGCATTTGTCTTGGAGAGGAA 2040
OY 2041 CAAGATTTCTTGAATTAATTTGGAAGAGGCTGCTATTAATGTTGTCAATAATGTTTT 2100
Db 2041 CAAGATTTCTTGAATTAATTTGGAAGAGGCTGCTATTAATGTTGTCAATAATGTTTT 2100
OY 2101 TGTAACTATGAAATTAATTAATCAGATGTAAATAGGAAAAAATTTTTTTTTTTTTT 2158
Db 2101 TGTAACTATGAAATTAATTAATCAGATGTAAATAGGAAAAAATTTTTTTTTTTTTT 2158
```

```
RESULT 2
; US-10-023-515-3
; Sequence 3, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-12201
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1746
```


TYPE: DNA
ORGANISM: Homo sapiens
US-10-023-515-3

Query Match 80.9%; Score 1746; DB 4; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

96 ATGCCAGAGGAGCTTACCTTCACTGCTTGAACATGAGTGTCTTTTCTGATTTCCGACGCC 155
1 ATGCCAGAGGAGCTTACCTTCACTGCTTGAACATGAGTGTCTTTTCTGATTTCCGACGCC 60
156 CTGTTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
61 CTGTTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
216 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTCTGGGAAGCCCTGTGCTGTG 275
121 ACAGAGCTGGATGATTCAGGGCAAGCAAGTCACTGTCTGGGAAGCCCTGTGCTGTG 180
276 AACGGTTCCTGGAGTCCCTTTGCTGCTCCCGGCTGGAGTCCCTGGATTTTAAAGAC 335
181 AACGGTTCCTGGAGTCCCTTTGCTGCTCCCGGCTGGAGTCCCTGGATTTTAAAGAC 240
336 CCGCAGCTGCATCGCCCTGGGATACTTGGAGAACCACTCTACCTTAATTTGTGC 395
241 CCGCAGCTGCATCGCCCTGGGATACTTGGAGAACCACTCTACCTTAATTTGTGC 300
396 CTCAGAACTCAGAGTGGCTGCTTTAGATCAACACATGCTCAAGGTGATTAACCGAAA 455
301 CTCAGAACTCAGAGTGGCTGCTTTAGATCAACACATGCTCAAGGTGATTAACCGAAA 360
456 TTCGAGATCAGAGACCTGCTCTTACCTTGAACATCTTATGCGCTCCGACCGGATA 515
361 TTCGAGATCAGAGACCTGCTCTTACCTTGAACATCTTATGCGCTCCGACCGGATA 420
516 GGCTCAAGCTCCCGCTCTTGGTGGTTCCGAGAGTGCCTTCAAGA CTGGCTCAGGC 575
421 GGCTCAAGCTCCCGCTCTTGGTGGTTCCGAGAGTGCCTTCAAGA CTGGCTCAGGC 480
576 TCATCTTTGATGGGTCCGCCCTGGCTGCTATGAGAGCTGTGTTGTCTGAC 635
481 TCATCTTTGATGGGTCCGCCCTGGCTGCTATGAGAGCTGTGTTGTCTGAC 540
636 TACCGGATGAGAAATTTGGTTCTTCAACAATGAGATCAAGTCTCTCGGGAACTGG 695
541 TACCGGATGAGAAATTTGGTTCTTCAACAATGAGATCAAGTCTCTCGGGAACTGG 600
696 GCTTCAAGGACAGGTGGCTGCTGATCTGGGTCCAGAGAACATCGAGTTCTTCGGT 755
601 GCTTCAAGGACAGGTGGCTGCTGATCTGGGTCCAGAGAACATCGAGTTCTTCGGT 660
756 GGGAGCCCGAGCTGTGTGACATCTTTGGAGTCCGCGGAGCCATAAGTGTCTAGT 815
661 GGGAGCCCGAGCTGTGTGACATCTTTGGAGTCCGCGGAGCCATAAGTGTCTAGT 720
816 CTTTATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 875
721 CTTTATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 780
876 GCATCATCTTAACTGAGAGGCCATGATTAAGAGAGTGAAGACCTGACAGTGGTT 935
781 GCATCATCTTAACTGAGAGGCCATGATTAAGAGAGTGAAGACCTGACAGTGGTT 840
936 GCAATTTCTGTGTAAATGAGTGTGAGACTGAGAGCCCTGTGAGTGTCTGAGAGCA 995
841 GCAATTTCTGTGTAAATGAGTGTGAGACTGAGAGCCCTGTGAGTGTCTGAGAGCA 900
996 AAAACCTTCAAGAGCTGTGAGACCTTCAAGCAGAAAAAAGTCTTCACTGAGTGGTT 1055
901 AAAACCTTCAAGAGCTGTGAGACCTTCAAGCAGAAAAAAGTCTTCACTGAGTGGTT 960
1056 GATGGTCTTTCTTCTTAATGAGCTCTAGATCTATTGCTCAGAAAGCATTTAAAGCA 1115

961 GATGGTCTTTCTTCTTAATGAGCTCTAGATCTATTGCTCAGAAAGCATTTAAAGCA 1020
1116 ATTCCTTCATCATGAGAGTCAATAACCAAGAGTGGCTTCTGTGCTTATGAAGAG 1175
1021 ATTCCTTCATCATGAGAGTCAATAACCAAGAGTGGCTTCTGTGCTTATGAAGAG 1080
1176 GCTCCAGAGTCTGAGTGGCTCCAAAGAGTCCCTGACCTGATGATTAACAAATC 1235
1081 GCTCCAGAGTCTGAGTGGCTCCAAAGAGTCCCTGACCTGATGATTAACAAATC 1140
1236 CTGACATCCCGCTCAGATTTTGAACCTTGTGGCTAATGAATTACTTCATGACAA 1295
1141 CTGACATCCCGCTCAGATTTTGAACCTTGTGGCTAATGAATTACTTCATGACAA 1200
1296 TCCCTGACTGAATCCGAGACAGTCTTCTGAGCTTGTGAGAGTGTCTTGTGTGTC 1355
1201 TCCCTGACTGAATCCGAGACAGTCTTCTGAGCTTGTGAGAGTGTCTTGTGTGTC 1260
1356 CCGCAGCTGATCAGAGCTGATATCAAGAGATGCTGGGACCTGTCTACTTCTATGAG 1415
1261 CCGCAGCTGATCAGAGCTGATATCAAGAGATGCTGGGACCTGTCTACTTCTATGAG 1320
1416 TTTGGCACCGGCTCAGTCTTTGAAGA CAGAAAGCCGCTTTGTCAAAGCCGACAC 1475
1321 TTTGGCACCGGCTCAGTCTTTGAAGA CAGAAAGCCGCTTTGTCAAAGCCGACAC 1380
1476 GCTGATGAAGTCCGCTTGTGTTCGGTGGTCTTCTGAAGGGGACATGTATGTTTC 1535
1381 GCTGATGAAGTCCGCTTGTGTTCGGTGGTCTTCTGAAGGGGACATGTATGTTTC 1440
1536 GAAGAGCCACGAGAGGAGAGAGTAACTGAGCCGAGATGAGTAATGAGGCTACC 1595
1441 GAAGAGCCACGAGAGGAGAGAGTAACTGAGCCGAGATGAGTAATGAGGCTACC 1500
1596 TTTGCTCGAACCGGAGATCTAATGGGAACGACTGTCTGTGGCAGCTTATATCTG 1655
1501 TTTGCTCGAACCGGAGATCTAATGGGAACGACTGTCTGTGGCAGCTTATATCTG 1560
1656 ACTGAGCAGTACCTCAGCTGGAATTGAACATGAGCTCTGGAACAGAGATCAAGAAC 1715
1561 ACTGAGCAGTACCTCAGCTGGAATTGAACATGAGCTCTGGAACAGAGATCAAGAAC 1620
1716 CCGGAGATTTTGAACAGACCAATCCCGGATGCTGCTGCGCCGACATGCTCAG 1775
1621 CCGGAGATTTTGAACAGACCAATCCCGGATGCTGCTGCGCCGACATGCTCAG 1680
1776 AGTCCCTTTCTTCTTAACTTCTCTCTCTCTCTCAAGCTTTCTTTCTTTGTGCT 1835
1681 AGTCCCTTTCTTCTTAACTTCTCTCTCTCTCTCTCAAGCTTTCTTTCTTTGTGCT 1740
1836 CCTTGA 1841
1741 CCTTGA 1746

RESULT 3
US-09-799-451-155
Sequence 155, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yunging
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong

/ APPLICANT: Wang, Zhiwei
 / APPLICANT: Wang, Dunhui
 / APPLICANT: Yang, Yonghong
 / APPLICANT: Wehrman, Tom
 / APPLICANT: Ghosh, Reena
 / APPLICANT: Drmanac, Radoje T.
 / TITLE OF INVENTION: No. 678369e1 Nucleic Acids and
 / FILE REFERENCE: Polypeptides
 / CURRENT APPLICATION NUMBER: 803
 / CURRENT FILING DATE: 2001-03-05
 / NUMBER OF SEQ ID NOS: 948
 / SOFTWARE: pc_fl_genes Version 2.0
 / SEQ ID NO 155
 / LENGTH: 965
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (128)..(925)
 / US-09-799-451-155

Query Match 23.6%; Score 508.6; DB 4; Length 965;
 Best Local Similarity 81.0%; Pred. No. 1.8e-141;
 Matches 675; Conservative 0; Mismatches 4; Indels 154; Gaps 1;

QY 186 GGGGCTTCTGCTGAGAGGCGCAGAGGAAACACAGGCTGGATGATTCAGGGCAAGCA 245
 DB 287 GGGGCTTCTGCTGAGAGGCGCAGAGGAAACACAGGCTGGATGATTCAGGGCAAGCA 346
 QY 246 GTACCTGTGCTGGAGAGGCGCTGCTGTGTAACGTGTTCTCTGGAGTCCCTTGTGCT 305
 DB 347 GTACCTGTGCTGGAGAGGCGCTGCTGTGTAACGTGTTCTCTGGAGTCCCTTGTGCT 406
 QY 306 CCCCCTGGGATCCCTGCAATTTAGAAACCGGACCTGCAATCGCCCTGGATTAAGT 365
 DB 407 CCCCCTGGGATCCCTGCAATTTAGAAACCGGACCTGCAATCGCCCTGGATTAAGT 466
 QY 366 CGAGAGGCACTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 425
 DB 467 CGAGAGGCACTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 526
 QY 426 CAACATGCTCAAGTGTGATTAACCGAATTCGAGTGTGAAAGATGCTGCTTAAGT 485
 DB 527 CAACATGCTCAAGTGTGATTAACCGAATTCGAGTGTGAAAGATGCTGCTTAAGT 586
 QY 486 AACATTTATGCGCTGCGCCAGCGCAATTCAGGCTCAAGCTCCCGCTTGGTGTTC 545
 DB 587 AACATTTATGCGCTGCGCCAGCGCAATTCAGGCTCAAGCTCCCGCTTGGTGTTC 646
 QY 546 CAGAGAGGCTTCAAGCTGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 605
 DB 647 CAGAGAGGCTTCAAGCTGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 706
 QY 606 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
 DB 707 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
 QY 666 ACATGGATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 725
 DB 766 ACATGGATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 785
 QY 726 TGGGTCCAGAGAAATCGAGTCTTCCGTGGGAGCCCAAGCTGTGTAACATTTTGGC 785
 DB 766 TGGGTCCAGAGAAATCGAGTCTTCCGTGGGAGCCCAAGCTGTGTAACATTTTGGC 785
 QY 786 GAGTCCGCGGAGGCAATGATTTTCTAATCTTCTCCATGAGCAAGGCTTA 845
 DB 766 GAGTCCGCGGAGGCAATGATTTTCTAATCTTCTCCATGAGCAAGGCTTA 845
 QY 846 TTCCCAAAAGCCATCATGAGAGAGTGGGTCATTCCTTCACTTGGAGGCTCATGAT 905
 DB 793 TTCCCAAAAGCCATCATGAGAGAGTGGGTCATTCCTTCACTTGGAGGCTCATGAT 852

QY 906 TATGAGAGAGTGAAGAGCTGCAAGTGTGCAATTTCTGTGTAAACAATGCTCAGAC 965
 DB 853 TATGAGAGAGTGAAGAGCTGCAAGTGTGCAATTTCTGTGTAAACAATGCTCAGAC 912
 QY 966 TCTGAGGCTCTGCTGAGTGTGCTGAGGACAAACCTCCAGAGGCTGCTAC 1018
 DB 913 TCTGAGGCTCTGCTGAGTGTGCTGAGGACAAACCTCCAGAGGCTGCTAC 965

RESULT 4
 US-09-949-016-3799
 / Sequence 3799, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / FILE REFERENCE: C1001307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / PRIOR FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / NUMBER OF SEQ ID NOS: 207012
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 3799
 / LENGTH: 2117
 / TYPE: DNA
 / ORGANISM: Human
 / US-09-949-016-3799

Query Match 18.8%; Score 406.6; DB 4; Length 2117;
 Best Local Similarity 56.5%; Pred. No. 1.2e-110;
 Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;

QY 211 GAAACCCAGGCTGGATGATTCAGGCGAAGCAAGTCACTGTGCGGAACCCGCTGC 270
 DB 157 GAAACCCAGGCTGGATGATTCAGGCGAAGCAAGTCACTGTGCGGAACCCGCTGC 216
 QY 271 CTGGAACGTTCTCTGGAAGTCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 330
 DB 217 GGTTCGAACCTTCTGGAAGTCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276
 QY 331 GAAACCCAGGCTGGATGATTCAGGCGAAGCAAGTCACTGTGCGGAACCCGCTGC 390
 DB 277 GAAACCCAGGCTGGATGATTCAGGCGAAGCAAGTCACTGTGCGGAACCCGCTGC 336
 QY 391 TGTGCTCAGAACTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 447
 DB 337 TGTGCTCAGAACTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
 QY 448 ACCCGAATTCGGAAGTCAAGAGCTGCTCACTCACTCACTCACTCACTCACTCACT 507
 DB 397 TCCCTTCGACCTCACTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456
 QY 508 CCGATTCAGGCTCAAGCTCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
 DB 457 GCGATTCAGGCTCAAGCTCCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
 QY 568 GCTCAGCTCACTTGTGATGGGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 627
 DB 517 GATGAGCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
 QY 628 TGTTCAGTACCGGCTGAGGAATATTGATTTCTTCAACAATGAGATCAGATGCTCGG 687
 DB 577 TGTTCAGTACCGGCTGAGGAATATTGATTTCTTCAACAATGAGATCAGATGCTCGG 636
 QY 688 GGAAGTGGGCTTCAAGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747

```

Db 637 GCAACTGGGGCTACCTGACCAAGTGGCTGCACTACGCTGGGTCCAGAGAAATATGCC 696
Qy 748 TCTTGGTGGGAGACCCGACGCTGTGACATCTTTGGCGAGTCCCGGGAGCATTAAG 807
Db 697 ACTTTGGAGGCAACCTTGACCGGTGTACATTTTGGCGAGTCTGGGGTGGACAGATG 756
Qy 808 TTTTCACTTATCTACTGTCTCCCATGCGCAAGGCTTATTTCCAAAGCCATCATAGAGA 867
Db 757 TGTCTTCGCTTGTGTGTCTCCCATATCCCAAGAGACTTTCACGAGAGCATATAGAGA 816
Qy 868 GTGGGGTGGCCATCATCTCTTAATCTGAGAGCCCATGATTAAGAAAGTGGAGACTGC 927
Db 817 GTGGGGTGGCCCTCTCTGCGCGGCTCATATGCGAGCTCAAGCTGA-----TGTATCTCA 870
Qy 928 AGTGTGTGCAATTTCTGTGTGAACAATGCGTCAAGCTCTGAAGCCCTGCTGAGTGGC 987
Db 871 CGGTGTGGCCCAACCTGTGTGCTGTGACCAAGTGAATCTGAAGCCCTGTGGGTGGC 930
Qy 988 TGAAGACAAACCTTCGAAGAGCTGTGACCCCTCAGCAGAAACAAAGTCTTTCATCTC 1047
Db 931 TCGGGGGCAAGATTAAGAGAGATTTCTTGCAATTAACAAGCTTTCAAGATGATCCCG 990
Qy 1048 GAGTGTGTGATGATGCTTCTTCTCTAATGAGCTTCAATCTAATGCTCAAGAAAGCAT 1107
Db 991 GAGTGTGTGATGAGGCTTCTCTGCGGAGCAACCCAGAGAGCTGTGGCTGTGCGCACT 1050
Qy 1108 TTAAGCAATTTCTTCCATCATGAGTGAATTAACAGAGTGGCTTCTGTGCTCA 1167
Db 1051 TTCAAGCTGTCTCTGACATGTGTGTGTCACAAACATGAATTCGGCTGTCTATCCCA 1110
Qy 1168 TGA-----AGAGGCTCTGAGATCTCTGAGTCTCAAGTCCCTTGCC 1215
Db 1111 AGGTCAATGAGATCTAATGATCCCAAGAGAAATGAAGAGAGGCTCCAGGCTGTCTC 1170
Qy 1216 TCCATCTAATACAAACATCTGACATCCGCTCAATTTTGAACCTTGTGTGAATG 1275
Db 1171 TCCAGAAATGTAAAGCTGTGATGTGCTCTCAATTTGTGATCTGTGAGAGAGAG 1230
Qy 1276 AATACTTCATGACAAAGCACTCCCTGACCTGAATTCAGAGAGAGTCTTGTGACTTGG 1335
Db 1231 AGTACATTTGGGAGCAATGGGAGATCCCAAGACCTTCAGAGAGAGTTCAGAGATGATG 1290
Qy 1336 GAGATGTGTCTTGTGTGCTCCCTGACCTGATCAACAGCTGATATCAAGAGATGCTGTG 1395
Db 1291 CGGATCTCAATGTTTGTGATCTCCCTGACCTCAAGTGAACATTTTCAAGT-----TTCGGGG 1347
Qy 1396 CACTGTCTACTTCTAATGAGTTCGGCAACGGGCTCAATGCTTGTGAAGACAGAGCCGG 1455
Db 1348 CCCGTGTACTTCTAATGAGTTCAGAGATCCAGCTGAGCTGTGCAAGAAATCAAGGCCAC 1407
Qy 1456 CTTTGTCAAGCCGACAGCTGATGAATCCGCTTGTGTGTGGTGTGCTTCTGA 1515
Db 1408 CCACATGAGAGGACAGCAATGATGAGACTTCTTGTTCAGAAATCTTGTGGGG 1467
Qy 1516 AGGGGAGCACTTGTATGTTGAAAGAGAGCAAGAGAGAGAAATTAAGTGAAGCCGAGAGA 1575
Db 1468 GCAACTCAATTAAT-----CACTGAAGAGAGAGAGCAAGCTTAAGAGAGAGAGA 1515
Qy 1576 TGATGAATATCTGGGCTACTTGTGCTGAACCGGAGAACTTAATGGAGAGAGAGCTGTCTC 1635
Db 1516 TGATGAATATCTGGGCTACTTGTGCTGAACCGGAGAACTTAATGGAGAGAGAGCTGTCTC 1575
Qy 1636 TGTGTGAGCTTATTAATCTGACTGAGAGTACTCTCAAGCTGGAATTTGAACATGAGCTGTG 1695
Db 1576 ACTGTGCGCTGTTCGACAGAGAGAGCAATACCTGCACTGAACCTTAAGAGAGCTGTGCTG 1635
Qy 1696 GACAGAGACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1746
Db 1636 GCGGGGCTGAAGGCCACAGAGCTCAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686

```

RESULT 5
US-09-949-016-555

```

; Sequence 555, Application US/09949016.
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 555
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-555

Query Match 18.8%; Score 406.6; DB 4; Length 2169;
Best Local Similarity 56.5%; Pred. No. 1,2e-110;
Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;

Qy 211 GGAACACAGGCTGGAGTGAATTCAGGGCAGACAGTACTGTCTGGGAGCCCTGTGC 270
Db 157 GGAACACACAGGCTGGAGTGAATTCAGGGCAGACAGTACTGTCTGGGAGCCCTGTGC 216
Qy 271 CTGTGAAGTGTCTCTGGAGTCCCTTGTGCTGTCTCCCGCTGGAGATCCCTGCAATTA 330
Db 217 GGGTCCAACTTCTCTGGAGATTCATTTGACCAAGCAGCTTAAGTCCGCTGCAATTTG 276
Qy 331 CGAACCCGAGCTGATCGCCCTGGAGTAACTTGCAGAGAGCCATCTCAATTT 390
Db 277 CACCCCTGAGCCCTCTGAATCTTGAAGTGTGAGGAGTGAACACCACTCCGAGCA 336
Qy 391 TGTGCTCCAGAACTCAGAGTGGCTCTTGA---TCAACATGCTCAAGTGCAT 447
Db 337 TGTGTCTACAGAGCTCAACCGAGTGAAGTCAAGTCTTGAACAGTTCAATGACT 396
Qy 448 ACCGGAATTTGGAAGTGAAGAGCTGCTTACTCTGAACATCTATGAGGCTGCCCCAG 507
Db 397 TCCCTTCCAGCTTCATGCTGAGAGTGTCTTACTCAGCATCTACAGCCGAGCCATA 456
Qy 508 CCGATACAGGCTCAAGCTCCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 567
Db 457 GCGATGAAGGCTTAACCTGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 516
Qy 568 GCTCAGCTCATCTTGAATGAGGCTCGCCCTGAGCTCTATAGAGAGCTGTGTGTGTG 627
Db 517 GCATGGCTCTCTTGAATGAGGCTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
Qy 628 TGTCTCAAGTACCGGCTAGAGAAATTTGTGTCTTCAACATGAGATCAAGATGCTCGG 687
Db 577 TATCAAGTACCGGCTAGAGAAATTTGTGTCTTCAACATGAGATCAAGATGCTCGG 636
Qy 688 GGAATCGGCTTCAAGAGCAGAGTGTGCTGTCTCTGAGTGTGTGTGTGTGTGTGTGTGT 747
Db 637 GGAATCGGCTTCAAGAGCAGAGTGTGCTGTCTCTGAGTGTGTGTGTGTGTGTGTGTGT 696
Qy 748 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 807
Db 697 ACTTTGAGGCAACCTTGAACCTGTGACATTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 756
Qy 808 TTTTCACTTATCTACTGTCTCCCATGCGCAAGGCTTATTTCCAAAGCCATCATAGAGA 867
Db 757 TGTCTTCGCTTGTGTGTCTCCCATATCCCAAGAGACTTTCACGAGAGCATATAGAGA 816
Qy 868 GTGGGGTGGCCATCATCTCTTAATCTGAGAGCCCATGATTAAGAAAGTGGAGACTGC 927

```

```

Db      817 GTGGCGTGGCCCTCTGCGCCGCGCTCATTTGCGAGCTGA-----TGTCACTTCCA 870
Qy      928 AGTGGTTGACATTTCTGTGTAAACAATGCTCAGACTCTGAGGCCCTGCTGAGTGGC 987
      871 CGGTGTGGCCCAACTGTCTGTGTGAACAATTGACTCTGAGGCCCTGTGTGGGCTGGC 930
Qy      988 TGAGGACAAAACCTTCAGAGGCTGTGACCTTCAGGACAAAACAAAGCTTTTCACTC 1047
      931 TCGGGGGCAGAGTAAGAGGAGATTCTTGCAATTAAACAACCTTTCAAGATATCCCG 990
Qy      1048 GAGGTGTGATGAGTCTTTCTTCTTAATGAGCCTTAATATGTTGTCTCAAGAACAT 1107
      991 GAGTGTGAGTGGGTCTTCTCTGCGCAGGCAACCCCAAGAGCTGTGGCTCTGCGACT 1050
Qy      1108 TTAAGCAATTCCTTCCATCATCGAGATCAATTAACAAGTGTGGCTTCTGTGCTTA 1167
      1051 TTCAAGCTGTCTTACGACTTGTGTGTCAACAATTAATTCGGCTGCTCATATCCCA 1110
Qy      1168 TGA-----AGAGGCTCTGAGATCTCAATGCTCCACAGTCCCTTGGCC 1215
      1111 AGGTCAATGAGATCTATGATATCCAGAAAGAAATGACAGAGAGGCTCCAGGCTGCTC 1170
Qy      1216 TCCATCATATCAAAAATCCTGCAATCCCGCTCAATTTTGACCTTGTGGCTAATG 1275
      1171 TGCAGAAATGTTAACTGTCTGTATGTTGCTCTTCACTTGTGTGACTGTGAGGAGG 1230
Qy      1276 AATTCCTTCAATGACAAAGCACTCCCTGAATGAAATCCGAGACAGTCTTGTGAACTTGG 1335
      1231 AGTACATTTGGGGCAATGGGGATCCCAAGACCTCCAGAGGCAATTCAGAGATATATG 1290
Qy      1336 GAGATGTGTTCTTGTGTGCTCCGCACTGATCAAGCTGATATCAAGATGTGTG 1395
      1291 CGGATCTCAATTTGTGTATCTTGTGCACTCAAGTACACATTTTCAAGT---TTCCCGG 1347
Qy      1396 CACTGTCTAATCTTCTAATGAGTTTCGACCGGCTCATGCTTGTGAAACACAGAGCGG 1455
      1348 CCGCTGTGATCTTCAAGAGTTCCAGATCAAGCCAGCTGTGCTCAAGAACTACAGCCAC 1407
Qy      1456 CTTTGTCAAGCCGCAACGCTGATGAAGTCCGCTTGTGTGTGGTGTGCTTCTGA 1515
      1408 CGCAATGAAGGAGACCAATGATGAGACTTCTTGTGTTTCAAGAGTTCTTTTGGGG 1467
Qy      1516 AGGGGACATTTGATGTTGAAAGGACCAAGAGAGAGAAATTAATGAGCCGAGAA 1575
      1468 GCAACTACATTAAT-----CACTGAGAGAGAGACAGCTTAAGAGAGAA 1515
Qy      1576 TGAATGAATACCTGGCTTACCTTGTCTGAACCGGAAATCTTAATGGAACGACTGTCTC 1635
      1516 TGAATGAATACCTGGCTTACCTTGTCTGAACCGGAAATGGAACCGGAGGTGTGACAC 1575
Qy      1636 TGTGACCAAGCTTATATCTGACTGAGAGTACTGCACTGAGACTTGAACATGAGCTG 1695
      1576 ACTGCGCCCTGTTCACACAGAGAGAGCAATACCTGCACTGAACCTTCAAGCTGCGGTG 1635
Qy      1696 GACAGAGACTCAAGAACCGCGGAGTATTTTGGACCAAGCAATCCCGC 1746
      1636 GCGGGCTCTGAAGGCCACAGAGCTCCAGTTCTGGAAGAGCGCTGCCCC 1686

```

RESULT 6

```

US-09-595-682B-27
; Sequence 27, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SU-0005
; CURRENT APPLICATION NUMBER: US/09/595, 682B
; PRIOR APPLICATION NUMBER: 60/075, 258

```

```

; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 2191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-595-682B-27

```

```

Query Match      18.8%; Score 406.6; DB 4; Length 2191;
Best Local Similarity 56.5%; Pred. No. 1,2e-110;
Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;

```

```

Qy      211 GGAACACCAAGCTGGAGATGATTCAGGGCAAGCAATCTGTCTGGAAAGCCCTGTGC 270
      179 GGAACACACACAGGGGAGGTGCTGGGAGTCTTGTCAATGTAAGAGGGCCCAATCCG 238
Qy      271 CTGTGAACGTGTTCTGTGGAGTCCCTTGTGCTGCCCGCTGGATCCCTGGATTTA 330
      239 GGGTTCAAACCTTCTGGGAATTCATTTGCAAGCAACCTTAAGTCCGCTGATTTG 298
Qy      331 GGAACCGGACGCTGCAATCCGCTGGAGTAACTTGCAGAGCCACTCTTACTTAATT 390
      299 CACCCCTGAGCCCTCTGAATCTTGTGAGTGTGAGAGATGAAACCAACCATCCGACCA 358
Qy      391 TGTGCTTCAGAACTCAAGTGTGCTGCTTAA-----TCAACATATGCTCAAGTGCATT 447
      359 TGTCTTACAGAGACTCACCGAGTGAATCAAGTTCTTACAGCTCAATATACCT 418
Qy      448 ACCGAAATTCAGAGTTCAGAAAGATGCTCTTAACTTAAATCTAATGCGCTGCGCAG 507
      419 TCCCTTCCAGCTCAATGTGAGAGACTGCTGTACTCAAGCACTTCAACGCGGCCATA 478
Qy      508 CCAATACAGGCTCAAGCTCCCGCTGTGTGTGTTCCAGAGAGGCTTCAAGACTG 567
      479 GCAATGAAGGCTTAACCTGCTGATGTGTGATCAAGGTGTGCTGCTTGTGTTTG 538
Qy      568 GCTACGCTCATCTTATGATGAGTCCGCTGTGCTCTTAATGAGACGCTGTGTGTG 627
      539 GCAATGCTCTTGTATGATGTTCCATGCTGCTCTTGAAGACGTTGTGTGTA 598
Qy      628 TCGTCAAGTACCGGCTAAGAAATTTGTTCTTACCAATGGATACGATGCTCCG 687
      599 TCAATCAAGTACCGCTGCTGCTGCTTCTTACCACTGAGACCAAGCAAGCAAGC 658
Qy      688 GGAACCTGGGCTTCAAGAACCAAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 747
      659 GCAACTGGGGCTACCTGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
Qy      748 TCTTGGGTGGGAGCCCAAGCTGTGACCATCTTGTGGAGTCCCGGAGCCATTAATG 807
      719 ACTTTGAGGAGCAACCTGACGCTGTACCAATTTTGGAGTGTGCGGTGTGAGCGAGT 778
Qy      808 TTTCTAGTCTTATCTGTCTTCCATGCGCAAGCTTATTCACAAAGCCATCAATGAGA 867
      779 TGTCTGCTGTGTGTGCTTCCCATATCCAAAGACTTTTCAAGAGCCATCAATGAGA 838
Qy      868 GTGGGTGGCCATCATCTCTTACCTGAGGCGCATGATTAATGAGAGAGAGCACTGC 927
      839 GTGGGTGGCCCTCTGCGCGGCTCATTTGCACTGAGCTGA-----TGTCACTTCCA 892
Qy      928 AGTGGTTGACATTTCTGTGTAAACAATGCTCAGACTGAGAGCCCTGTGAGTGGC 987
      893 CGGTGTGGCCAACTGTCTGTGTGAACAATGACTGTGAGAGCCCTGTGTGGCTGGC 952
Qy      988 TGAGACAAAACCTTCAGAGGCTGTGACCTTCAAGCAAGAAACAAAGTCTTCACTC 1047
      953 TCGGGGCAAGAGTAAGAGAGATTTGCAATTAACAAGCTTTCAAGATATCCCG 1012
Qy      1048 GAGTGTGATGATGCTTTCTTCTAATGAGCTTAAATGATTTGTCTCAAGAAAGCAT 1107

```

Db 1013 GAGTGTGATGAGGGCTCTTCCCTGCCAGGCAACCCAGAGAGCTGCGCTCTGCGCACT 1072
 Qy 1108 TTTAAGCAATTCCTTCATTCAGAGTCAATTAACAGAGTGTGCTTCCGCTGCTA 1167
 Db 1073 TTGAGCCCTGCTCCTGAGATGTGTGTTCACAAACAAATTAATTCGCTGCTCATCCCA 1132
 Qy 1168 TGA-----AGGAGGCTCCTGAGATCCTGAGTGTCTCAACAAGTCCCTTGGCC 1215
 Db 1133 AGGTATGAGGATCTATGATACCCAGAGAGAAATGAGACAGAGAGGCTCCAGAGGTGCTC 1192
 Qy 1216 TCCATCTGATACAAAACATCTGCAATCCGCTCAAGTATTGCACTTGTGCTTAATG 1275
 Db 1193 TGCAGAAATGTAAAGCTGCTGATGTGCTCTCAATTTGTGTGACCTGCTGAGAGGAG 1252
 Qy 1276 AATCTTCATGACAGAGCACTCCCTGACTGAAATCGAGACAGTCTTCTGACTTGTG 1335
 Db 1253 AGTACATTTGGGAGCAATGAGGATCCCAAGACCTTCAGAGCGAGTTTCAAGAGATGATGG 1312
 Qy 1336 GAGATGTGTCTTGTGTGCTCCTGCACTGATCAAGCTGATATGACAGAGATGCTGGTG 1395
 Db 1313 CGGACTCATGTGTGTGATCCCTGCACTCAAGTACGACATTTTCAAGT---TTCGGGG 1369
 Qy 1396 CACCTGTCTACTTCTATGATGTTTGGGCAACGGCTCAAGTCTTTGAAGACAGAGCCGG 1455
 Db 1370 CCCCTGTACTTCTATGAGTTCCAGATTCAGCCAGCTGGCTCAAGAAATCAAGGCCAC 1429
 Qy 1456 CTTTGTGTAAGCCGACACGCTGATGAAATCCGCTTGTGTGCTGTGTGCTTCTGGA 1515
 Db 1430 CGCAGATGAGGCAACAGATGATGAGCTTCTTGTGTGTAAGATTTCTTTGGGG 1489
 Qy 1516 AGGGGGAATTTGTTATGTTTCGAGAGGACGAGAGAGAGAAATTAATGAGCCGGAAGA 1575
 Db 1490 GCACTACATTTAAAT-----CACTGAGAGAGAGAGAGAGAGTAAAGAGAGAA 1537
 Qy 1576 TGATGAAATCTGAGCTACTCTTGTCTGAAACGGGAATCTTAATGGAACGACTGTCTC 1635
 Db 1538 TGATGAAATCTGAGCTACTCTTGTCTGAAAGAGAGAGAGAGAGAGTAAAGAGAGAA 1597
 Qy 1636 TGTGCGCAGCTTATATCTGACTGAGAGTACTCTGACTGAGACTTGAAGAGAGAGCTG 1695
 Db 1598 ACTGCGCCTCTGTCAGACGAGAGAGCAATCTGCACTGAACTTAAGAGAGAGCTG 1657
 Qy 1696 GACAGAGACTCAAGAGACGCGGCTGATGATTTTGGACCGAGCAATCCGCC 1746
 Db 1658 GCCGGGCTCTGAAGGCCACAGAGCTCAAGTTCTGAAAGAGAGAGCTGCCCC 1708
 RESULT 7
 US-09-595-682B-20
 ; Sequence 20, Application US/09595682B
 ; Patent No. 6800483
 ; GENERAL INFORMATION:
 ; APPLICANT: Danks, Mary K.
 ; APPLICANT: Potter, Philip M.
 ; APPLICANT: Houghton, Peter J.
 ; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
 ; Tumor Cells
 ; FILE REFERENCE: SJ-0005
 ; CURRENT APPLICATION NUMBER: US/09/595,682B
 ; PRIOR FILING DATE: 2000-01-16
 ; PRIOR APPLICATION NUMBER: 60/075,258
 ; PRIOR FILING DATE: 1998-02-19
 ; PRIOR APPLICATION NUMBER: PCT/US99/03171
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 20
 ; LENGTH: 1717
 ; TYPE: DNA
 ; ORGANISM: Oryctolagus cuniculus
 US-09-595-682B-20
 Query Match 16.8%; Score 362.6; DB 4; Length 1717;

Best Local Similarity 54.8%; Pred. No. 1,66-97;
 Matches 880; Conservative 0; Mismatches 649; Indels 78; Gaps 5;
 Qy 186 GGGCTTCTGCTGAGAGGCGCACAGAGAAACACAGCTGGATGATTCAGAGCAAGCAA 245
 Db 62 GGGCACCGGTCTGCAACACCTGTGTAGATCTGTCAATGGAAGTCTTGGGAAATTC 121
 Qy 246 GTCACTGTGAGGAAGCCCTGTGCTGTGAAAGTTCCTCGAGTCCCTTGTGCTCT 305
 Db 122 GTCACTGTGAGGAATTTGTCACAGCCCGTGTGCTTCTGTGAGTCCCTTGTGCT 181
 Qy 306 CCCCCGTGGAGTCCCTGATTTAGAACCCGAGCTGATGCTCCCTGGATTAATTTG 365
 Db 182 CCCCCCTTGGATCCCTGAGTTTGTGACACACAGCTGTGAGATCAATGAGCCAGTG 241
 Qy 366 CGAGAGCCACTCTTACCTTAATTTGTGCTCCAGAAC-----TCAGAGTGGTGTG 419
 Db 242 AAGAACACACCTCTTACCTTCCATGTGCTCCAGAGCGAGATTCAGGGCATATGCTC 301
 Qy 420 TTAGATCAACATGCTCAAGGTGATTAACCCGAAATTCGAGTGTCAAGAGCTGCTC 479
 Db 302 TCGAGCTCTTCAACACAGAAAGAAACATCCCTTTAATTTTCTGAGAGCTGCTT 361
 Qy 480 TAACCTGAACATCTATGCTGCTGCCAGCGAGATACAGGCTCCAGCTCCCGTCTTGTG 539
 Db 362 TACCTGAATATTACACCCCTGTGACCTGAACAAGAGAGAGAGCTCCGCTGATGGTG 421
 Qy 540 TGTGTTCCAGAGAGTCTTCAAGACTGTGCTGACCTCAATTTTGAATGGTCCGCTG 599
 Db 422 TGGATTCATGAGAGGTGTCTATGTTGTGTGAGAGATCAACCTTAATGATGGCTGT 481
 Qy 600 GCTGCTTATGAGAGAGTGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659
 Db 482 TCTGCCATGAGAGAGTGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
 Qy 660 TTCAACATGAGAGATCAAGCTGCTGCGGAAATCTGAGCTTCAAGAGCAAGTGTGCT 719
 Db 542 TTCAGCAGAGAGATGAGCAGCGGAGAGAACTGGGCTCACTTGAGCAGGTTGCTGG 601
 Qy 720 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
 Db 602 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
 Qy 780 TTTGGCAGTCCGCGGAGCAATAGTGTCTTATGCTTATGCTTATGCTTATGCTTAT 839
 Db 662 TTTGAGAGTCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 721
 Qy 840 GGGTTATTCACAAAGCCATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 899
 Db 722 AATCTCTTCAATGAGCAATTTCCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 781
 Qy 900 CATGATTAAGAGAGAGT 959
 Db 782 AAGAACACCAAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 841
 Qy 960 TCAAGCTGAGGCGCTGTGAGGCTGTGAGCAAAACCTCCAGAGAGCTGTGAC- 1018
 Db 842 TCGGCTGATGTGTCTACTGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 901
 Qy 1019 -----CTTACGACCAAAACAAAGTCT 1040
 Db 902 TTGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 961
 Qy 1041 TTCACTGAGT 1100
 Db 962 CTGACCACTGT 1021
 Qy 1101 AAAGATTTAAGCAATCTTCAATCAATGAGATCAATCAAGAGTGTGTGTGTGTGTGT 1160
 Db 1022 AAGAAATTCACATGCTGCTTCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1081
 Qy 1161 CTGCTTATGAGAGAGCTCTGAGATCTCTCAAGTGTCTCAACAAATGCTCTCTCTCT 1209

Db 1082 ATCCCAATGCAATGCTGGGCTATCCACTCTCTGAAGGCAAACTGACCAAGACAGCT 1141
 QY 1210 -TTGCCCTCCATCTGATACAAAACATCTCGACATCCCGCTCACTATTTGCACTTGTG 1268
 Db 1142 ACAGAACTCTTGTGGAAAGTCTTACCCATCTGCAATGCTCTAAAGAGCTGATCCAGTG 1201
 QY 1269 GCTAATG---ATACTTCATAGCAAGACATCCCTGACTGAATTCGAGAGAGCTTCTG 1325
 Db 1202 GCCACTGAGAGATTTTGAAGAGGACAGATGACCTCTGCAAAAAGAAAGACTTGTCTTG 1261
 QY 1326 GACTTGTCTGAGATGTGTTCTTGTGTGCTCCCTGACTGATCAGAGCTGATATACAGA 1385
 Db 1262 GACATGCTTGAAGATTGTATTGTGTGTCCATCTGAAATGTGGCTCGTACCAAGA 1321
 QY 1386 GATCTGTGACCTGCTCTACTTCTGATGAGTTTGGGACCGGCTCACTGCTTTGAAAGAC 1445
 Db 1322 GATCTGTGAGCCCACTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1381
 QY 1446 ACAGAGCCGCTTTTGTCAAGCCGACACGCTGATGATGATGATGATGATGATGATGATGAT 1505
 Db 1382 ATGAGACCCAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441
 QY 1506 GCTTCTGAAAGGGGACATTTGTATGTTGAAAGAGCCACGAGAGAGAGATTAAGT 1565
 Db 1442 CCGTTTAA-----AGAGGGTCCACAGAGAGAGATCAAACTG 1483
 QY 1566 AGCCGGAATGATGAAATATGCTGCTACCTTGTCTGAAACCGGAAATCTTAATGGAGAC 1625
 Db 1484 AGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1543
 QY 1626 GACTTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1685
 Db 1544 GGGCTTCTCAATGAGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1603
 QY 1686 ATGAGCCTGACAGAGACTCAAGAACCGGCGGTGATTTTGTGAC 1732
 Db 1604 ACCAGAGAGCCAGAACTGAAGACAGAGAGATGCTTTCTGAC 1650

RESULT 8

US-09-264-737-3
 / Sequence 3, Application US/09264737A
 / Patent No. 6107549
 / GENERAL INFORMATION:
 / APPLICANT: Feng, Paul C.C.
 / APPLICANT: Ruff, Thomas G.
 / TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
 / FILE REFERENCE: 38-21(10551) RLB3 Pyridine Tolerance
 / CURRENT APPLICATION NUMBER: US/09/264,737A
 / EARLIER FILING DATE: 1999-03-09
 / EARLIER APPLICATION NUMBER: 60/077,377
 / NUMBER OF SEQ ID NOS: 11
 / SOFTWARE: Patent Ver. 2.0
 / SEQ ID NO 3
 / LENGTH: 1701
 / TYPE: DNA
 / ORGANISM: Rabbit
 US-09-264-737-3

Query Match 16.5%, Score 357, DB 3, Length 1701,
 Best Local Similarity 54.8%, Pred. No. 7,5e-96,
 Matches 866, Conservative 0, Mismatches 635, Indels 78, Gaps 5,
 QY 214 ACACGAGCTGGGATGATTTAGGCGAAGCAAGTCACTGTGCGAAGCCGTGTGCTG 273
 Db 83 ACACCTGAAAGGGAAGTCTGTGGGAAAGTTCGTCACTTAGAAGATTGACAGCCG 142
 QY 274 TGAACGTGTTCTCGAGTCCCTTTGTGCTGCTCCCGCGGAGATCCCGCAATTACGA 333
 Db 143 TGGCGTCTTCTCGAGTCCCTTGTGCTGCTCCCGCGGAGATCCCGCAATTACGA 202

QY 334 ACCGAGGCTGATGCGCCGTGGATTAATTGCGAAGCCACTTCTTAACCTTAATTGT 393
 Db 203 CACCAAGCTCTGATTAATGAGACACGTGAAGAACCACTCTTAACCTTCCATGT 262
 QY 394 GCTTCAAGAC-----TCAGATGCTGCTTTAGATCAACACATGCTCAAGTGCA 447
 Db 263 GCTTCCAGAGCCAGATATACAGGCAATATGCTCTCGAGCTTCAACCAAGAAAGAG 322
 QY 448 ACCGAAATTTGAGATGTCAGAGATGCTGCTTCTACTTAACAATCTATGCGCTGCCAG 507
 Db 323 ACATCCCTTTAAGTTTCTGAAGATGCTCTTACCTGAATATTTACACCTCTGAC 382
 QY 508 CCGATPACAGCTCCAAAGCTCCCGCTTGTGTGTGTGCTCCAGAGAGTCCCTTCAAG 567
 Db 383 TGAACAAAGAGAGGAGGCTGCGGATGATGATGATGATGATGATGATGATGATGATGAT 442
 QY 568 GCTGAGCTTCACTTTATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
 Db 443 GTGAGCATCAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502
 QY 628 TGTCTCACTACCGGCTAGAAATATTTGATTTCTTCAACACATGAGATACATGCTCCG 687
 Db 503 CCATTCAGTACCGGCTGAGATCTGGGATTTCTTACACACAGAGATGACACAGCCGAG 562
 QY 688 GGAATGAGCTTTCAAGACCAAGTGTGCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 747
 Db 563 GGAATGAGCTTTCAAGACCAAGTGTGCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 622
 QY 748 TCTTGTGAGGAGACCCAGGCTGTGACCACTTTTGGCCAGTCCGCGGAGCCATAGT 807
 Db 623 ACTTTGAGAGGAGACCCAGGCTGTGACCACTTTTGGCCAGTCCGCGGAGCCATAGT 682
 QY 808 TTTTATGCTTTATAGTGTCTCCATGCGCCAAAGCTTATTTCAAGACATCATAGAGA 867
 Db 683 TCTTATCTTTATATATCTCCCTGACCAAGATCTTTTCAATGACCAATTTTCCAGA 742
 QY 868 GTGAGGAGGACCATATCCCTTACTGAGAGCCATATATGAAAGATGAGAGCTTGC 927
 Db 743 GTGAGGAGGACCATATCCCTTACTGAGAGCCATATATGAAAGATGAGAGCTTGC 802
 QY 928 AGGTGTTGACATTTCTGTGTAAATGCTGACATCTGAGGCGCTGAGAGTGC 987
 Db 803 TTGCAATGCAAGCTGTGTGTAACCAACCACTGCTGCTGCTGCTGCTGCTGCTGCT 862
 QY 988 TGAAGCAAAACCTTCAAGAGCTGTGAC----- 1018
 Db 863 AGAAGCAAGAGAAAGATCATGAGGTGACATTTGAATGAATTTATGCTTATGATC 922
 QY 1019 -----CCTAGCCAGAAACAAAGTCTTCTGCTGAGTGTGATGCTTCT 1068
 Db 923 TAGTTGGGACCCCAAGAGAACACCGCTTCTGACATGATGATGATGATGATGATGATGAT 982
 QY 1069 TTCTTATGAGCTTCAATGATTTATGTTGTGAGAAAGATTTAAAGCAATTCCTTCATCA 1128
 Db 983 TGCCAAAAGCACTTCAAGATTTCTGCAAGAAAGAAATACAAATGTGCTGCTCAATG 1042
 QY 1129 TCGAGTCAATTAACAGAGTGTGCTTCTGCTGCTGCTTATGAAGAGCTCTTGAATTC 1188
 Db 1043 TGGGATCAACAGGAAGTGTGCTGATTTATTCATGCAATGCTGAGCTATTCAC 1102
 QY 1189 TCAATGCTCCAAAGTCTT-----GCCCTCATCTGATATCAAAACATTC 1236
 Db 1103 TCTTGAAGGCAATGAGACCAAGAACAGCTTACAAATCTTTGTGAAGTCTTACCCA 1162
 QY 1237 TGCATATCCCGCTCAATTTTGAACCTTGTGTATG-----AATCTTCAATGACAGC 1293
 Db 1163 TTGTCAATGCTCTTAAGAGCTGCTCAAGTGGCACTGAGAAATATTTAGAGGAGCAG 1222
 QY 1294 ACTTCTGATGAAATCGAGACAGCTTCTGAGCTTGTGAGATGATGATGATGATGATGATGAT 1353
 Db 1223 ATGACCTGCTCAAAAGAAAGATTTTCTGAGACATGCTTGTGAGATTTGTATTGTG 1282
 QY 1354 TCCCTGACATGATCAAGCTGATATCAAGAGATGCTGTGACCTGTCTTATG 1413

Db 1283 TCCTCATCTGTGATGATGCTCCGTCACACAGAGATCTGAGCCGCCACCTATATGTATG 1342
Qy 1414 ACTTTGCGACACCGGCTCAGTGTCTTGAAGACACGAGCCGCTTTTGTCAAGCCGAC 1473
Db 1343 AGATATCGGATGATGCGCCAGACCTTCTCATGACATGAGCCCAAGACAGATAGGGGACC 1402
Qy 1474 AGCTGATGAATGCTGCTTGT 1533
Db 1403 ATGGAGATGAGATCTTCTCTGTCTTAGAGCCCGCTTTTAAAGG----- 1449
Qy 1534 TCGAAGAGCCACGAG 1593
Db 1450 -----GATCCACAG 1504
Qy 1594 CTTTGTCTGACACCGGATCTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1653
Db 1505 ACTTGTCTGAGATGAG 1564
Qy 1654 TGAATGAGAGATGCTCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1713
Db 1565 ACAGAGAGAGATGCTCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1624
Qy 1714 CGCGGATGAGATTTTGAC 1732
Db 1625 AGGAGTGTGCTTCTGAC 1643

RESULT 9

US-09-949-016-13807/c
; Sequence 13807, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMERPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13807
; LENGTH: 168971
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(168971)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13807

Query Match 8.2%; Score 176; DB 4; Length 168971;

Best Local Similarity 100.0%; Pred. No. 3.7e-40;

Matches 176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 AAAACAGGCGCTGAG 71
Db 2482 AAAACAGGCGCTGAG 2423
Qy 72 TTCTTAATTAAGTACGATCCCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 131
Db 2422 TTCTTAATTAAGTACGATCCCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2363
Qy 132 TGTCTTTTCTGATTTCTCAAGCCCTGTTGGAGACACAGAGAGAGAGAGAGAGAGAGAG 187
Db 2362 TGTCTTTTCTGATTTCTCAAGCCCTGTTGGAGACACAGAGAGAGAGAGAGAGAGAGAG 2307

RESULT 10

US-09-810-861B-3

; Sequence 3, Application US/09810861B

; Patent No. 6770799

; GENERAL INFORMATION:

; APPLICANT: Mor, Tsafir S.

; APPLICANT: Soreq, Hermona

; APPLICANT: Arntzen, Charles J.

; APPLICANT: Mason, Hugh S.

; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN

; FILE REFERENCE: BTI-45

; CURRENT APPLICATION NUMBER: US/09/810,861B

; PRIOR FILING DATE: 2001-03-16

; PRIOR FILING DATE: 2000-03-17

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 3.1

; SEQ ID NO 3

; LENGTH: 5767

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: plasmid vector

; OTHER INFORMATION: pTM034.

US-09-810-861B-3

Query Match 6.5%; Score 139.6; DB 4; Length 5767;

Best Local Similarity 51.8%; Pred. No. 2.6e-30;

Matches 422; Conservative 0; Mismatches 374; Indels 18; Gaps 4;

Qy 75 TTAATTAAGTACGATCCCAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 134
Db 823 TTACAGAGATAGCCATGAGCTCCCGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 882
Qy 135 TTTTCTGATTTCTCAGCCCTGTTGGAGACACAGAGAGAGAGAGAGAGAGAGAGAGAG 194
Db 883 CCATCTCTTCT 942
Qy 195 GCTGAAGGCGCACAG 254
Db 943 GATGAGAGAGCTGT 1002
Qy 255 CTGGAG 314
Db 1003 CCGGAG 1056
Qy 315 GGATCCCTGCGATTTACGAG 374
Db 1057 GGAAG 1116
Qy 375 ACCTCTACCTTAATTTGTGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 434
Db 1117 ACAACCTTCAAG 1176
Qy 435 CTGAAGTGTGATTAACCGAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 494
Db 1177 ACCGAG 1236
Qy 495 GGCCTGCGCAG 554
Db 1237 ACACCATACCCCGGCTTACATCCCCA---CCCTGTCTCTGTGTGTGTGTGTGTGTGT 1293
Qy 555 GCTTTCAAG 608
Db 1294 GGCCTTCAAG 1353
Qy 609 GAG 665
Db 1354 GAG 1413
Qy 666 ACATGAGATACAG 725

```

Db      1414  CCGGGGAGCCGAGAGCCCGCCGCAATGAGTCTCTCGAATCAGAGGCTGCGCCCTGAG 1473
Qy      726   TGGGTCGAGAGAAACAATGAGTCTCTCGAGTGGAGACCCAGCTGTGATCCATCTTTGAC 785
Db      1474   TGGGTGAGAGAGAAAGTGGCAGCCCTTCGGGGGTGACCCGACATGATGACCTGTTTGG 1513
Qy      786   GAGTCGCGGAGAGCCATTAAGTGTCTTCTAATCTGTCTTCCATGAGCCAAAGCTTAA 845
Db      1534   GAGAGCCGCGGAGAGCCGCTCGGTGGGCAATGACCTGTCTCCGACCGCCGCGGCGCTG 1593
Qy      846   TTCCACAAAGCCATCATGAGAGAGTGGGGTGGCCA 879
Db      1594   TTCCACAGGCGCGTGTCTGACAGCGGTGCCCA 1627

```

RESULT 11

```

US-09-810-861B-4
; Sequence 4, Application US/09810861B
; Patent No. 6770799
; GENERAL INFORMATION:
; APPLICANT: Mor, Teafir S.
; APPLICANT: Soreq, Hermona
; APPLICANT: Amnzen, Charles J.
; APPLICANT: Mason, Hugh S.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN
; FILE REFERENCE: BTI-45
; CURRENT APPLICATION NUMBER: US/09/810,861B
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/190,440
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 4
; LENGTH: 14446
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (11862)..(12157)
; OTHER INFORMATION: Description of Artificial Sequence: plasmid vector
; OTHER INFORMATION: pTM036. Identity of sequence residues 11862-12157 unknown.
US-09-810-861B-4

```

```

Query Match      6.5%; Score 139.6; DB 4; Length 14446;
Best Local Similarity 51.8%; Pred. No. 5, 1e-30;
Matches 422; Conservative 0; Mismatches 374; Indels 18; Gaps 4;

```

```

Qy      75   TTAATAAAGTACGATCCCATGCGACAGAGGACTTACTTCAATCTGCTTCACAATGCTGC 134
Db      127   TTACGACGATAGCCATGAGCTTCCCGAGATGCTCTGTCACACGCTTCCCTGAGCTTCC 186
Qy      135   TTTTCTCGATTTCTGACGCCCCCTGTTGGACACAGACAGTGGGAAAAAATCTGGGCTTTCT 194
Db      187   CCATCTCTTCTCTCTCTCTCTCTGCTCTGAGTGAAGAGTGGGGCTGAGAGGCGCGGAG 246
Qy      195   GCTGAAGGGCCACAGAGGAAACACAGAGCTGGAGTGAATTCAGGCGAAGCAAGTCACTGTG 254
Db      247   GATGACAGCTGCTGTGATACGATGCGTGGGGCGGCTGCGGGGCATTTCCCTGAAGACC 306
Qy      255   CTGGAGAGCCCTGTGCTGTGATGACGTGTTCTCTGAGAGTCCCTTGTGCTCCCGCGCTG 314
Db      307   CCGCGGGGCGCTGTCTCT-----GTTTCTGTGGCATCCCTTTTGGAGCCACCATG 360
Qy      315   GGATCCCTGCGATTTTGAAGACCGGAGCTGCAATGCTCTGGAGTAACTTGGAGAGCC 374
Db      361   GGACCCCGTGTCTTCTGCGACCGGAGCCCAAGCAGCTTGTGTAGAGGGGTGATAGAGCTG 420
Qy      375   ACCTCTACCTTAATTTGTGCTTCAGAGTCAAGAGTGTGCTCTTAATTAATCAACATG 434
Db      421   ACAACCTTCCAGAGTGTCTGTACCAATATGTGACACCTTATTAACCAAGGTTTGAAGG 480

```

```

Qy      435   CTCAGTGCATTAACCCGAATTTCGAGTGTCAAGAGATGCTCTTACTTAACTTAAT 494
Db      481   ACCGAGATGTGAACCCCAACCGTATGCTGAGCGAGATGCTGTATCTTAACGCTGTG 540
Qy      495   GCGCTGCGCCACGCGGATTAAGGCTTCAAGCTCCCGCTCTTGATGTGTTCCAGAGGT 554
Db      541   ACAACCAATACCCCGGCTTAATATCCCA---CCCTGTCTGTGTGATTTATGGGGGT 597
Qy      555   GCGTTCAAGATGAGTCAAGCTC-----CATTTTGAATGAGTCCGCTGAGCTTAA 608
Db      598   GCGTTTACAGTGGGGCTCTCTCTTGAACGTGTACATGACGCTCCCTTCTGTACAGGC 657
Qy      609   GAGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
Db      658   GAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 717
Qy      666   ACATGAGATCAAGATCTCTCGGGAATGAGGCTTCAAGACAGATGAGTGTGTGTGTGT 725
Db      718   CCGGGAGACCGAGAGGCGCCGCGCAATGTGGGTCTCTGATCAGAGGCTGGCCCTG 777
Qy      726   TGGTTCAGAGAAACATGAGTCTCTGATGGGAGACCCAGCTGTGTGACATTTTGGC 785
Db      778   TGGGTGACAGAGAAACGTGGCAGCTTGGGGGTGACCGACATCAATGAGCTGTTTGG 837
Qy      786   GAGTCGCGGAGAGCCATTAAGTGTCTTGAATCTGTCTTCCATGAGCCAAAGCTTA 845
Db      838   GAGAGCGCGGAGAGCGCTCGGTGGGATGATGACCTGTCTCTCCGCGCCGCGGCGCTG 897
Qy      846   TTCCACAAAGCCATCATGAGAGTGGGGTGGCCA 879
Db      898   TTCCACAGGCGCGTGTCTGACAGAGGTGCCCA 931

```

RESULT 12

```

US-09-810-861B-5
; Sequence 5, Application US/09810861B
; Patent No. 6770799
; GENERAL INFORMATION:
; APPLICANT: Mor, Teafir S.
; APPLICANT: Soreq, Hermona
; APPLICANT: Amnzen, Charles J.
; APPLICANT: Mason, Hugh S.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN
; FILE REFERENCE: BTI-45
; CURRENT APPLICATION NUMBER: US/09/810,861B
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/190,440
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 5
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (11862)..(12157)
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: human acetylcholinesterase gene optimized for
; OTHER INFORMATION: expression in plants
US-09-810-861B-5

```

```

Query Match      6.5%; Score 139.4; DB 4; Length 1725;
Best Local Similarity 55.1%; Pred. No. 1, 2e-30;
Matches 341; Conservative 0; Mismatches 286; Indels 12; Gaps 3;

```

```

Qy      270   CCTGTGAACGTGTTCTCTGAGTCCCTTGTGCTCTCCCGCTGGAGTCTTGCATTT 329
Db      175   CCGTGTCTGTCTTCTCTGGGATCCCTTTGGAGGACACCAATGAGACCCGCTGCTTT 234
Qy      330   AGCAACCGGAGCGCTGATGCGCTGGAGTAACTTGGAGAGGACACTTCACTTAAT 389
Db      225   CTGCACCGGAGCCCAAGCAGCTTGTGTCAAGGGTGTATAGCTTACAACTTCAAGAT 294

```


COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06106
FILING DATE: 19920722
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPM/EAB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
PCT-US92-06106-1

Query Match 6.5%; Score 139.4; DB 5; Length 1845;
Best Local Similarity 55.1%; Pred. No. 1.3e-30;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

QY 270 CCTGTGAACGTGTTCTCTGGAGTCCCTTTGCTGCTCCCGCGTGGATCCCGCATTT 329
Db 175 CCGTCTCTGCTTCTCTGGAGTCCCTTTGCTGCTCCCGCGTGGATCCCGCATTT 234
QY 330 ACGAACCCGAGCGCTGCATCGCCCTGGAGTAATTGCGAAGCCACTCTCAACCTTAAT 389
Db 235 CTGCCACCGAGCCCAAGACGCTTGTGTAAGGGGTGTAGACGCTCAACCTTTCAGAGT 294
QY 390 TTGTGCTTCAGAACTCAGAGTGGCTGCTTTAGATCAACAATGCTCAAGGTGATTAAC 449
Db 295 GTCTGTACCAATATGTGACACCCCTAACCAGGTTTGAAGGACCGAGATGTGAAC 354
QY 450 CCGAATTCGAGAGTGCAGAACTGCTCTTACCTGAACATCTATGCCCTGCCAGCC 509
Db 355 CCGAACCGTGAAGTGAAGAGTGGCTGCTTACCTGAACATCTATGCCCTGCCAGCC 509
QY 510 GATPACAGCTCCAGCTCCCGCTTGTGAGTGTCCAGAGAGTGGCTTCAAGACTGGC 569
Db 415 CCTACATCCCCA---CCCTGTCTCTGTGTGATCTATGGGGGTGCTTTCAGTGGG 471
QY 570 TCAGCTC-----CATCTTTATGGGTCCGCTGCTGCTGCTTGAAGAGCTGTGCT 623
Db 472 GCTTCTCTTGAAGTGTGATGATGCGCTTCTTGTGTAACAAGCGAGAGACTGTGCT 531
QY 624 GTGTGCTGCAAGTACCGGCTAGAAATATTTGATTTCTTCAAC---ACATGGGATCAAGAT 680
Db 532 GTGTGCAATGAATACCGGCTAGAAATATTTGATTTCTTCAAC---ACATGGGATCAAGAT 680
QY 681 GCTCCGGGAACTGGGCTTCAAGAGCAGAGTGGCTGTGTCTGTGCTGCTGCTGCTGCTGCT 740
Db 592 GCTCCGGGAACTGGGCTTCAAGAGCAGAGTGGCTGTGTCTGTGCTGCTGCTGCTGCTGCT 651
QY 741 ATCAGATTTCTCGGTGGGAGCCCAAGCTGTGACATCTTTTGGCAGGTCCGGGGAGCC 800
Db 652 GTGGACGCTTCCGGGGTGAACCCGATCAAGTACGCTGTTTGGGAGAACCGGGAGGCC 711
QY 801 ATAAAGTTTCTAATCTTAAATCTGTCTCCATGGCCAAAGCTTATTTCCAAAGCATC 860
Db 712 GCTCTGGGTGAGCACTGTGTGTCTCCGCGCCAGCGGGGCTGTTCACAGAGGCTGTG 771

QY 861 ATGAGAGTGGGAGTGGCCA 879
Db 772 CTGACAGCGGTGCCCCCA 790

RESULT 15
US-09-949-016-1192
Sequence 1192, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 1192
LENGTH: 2158
TYPE: DNA
ORGANISM: Human
US-09-949-016-1192

Query Match 6.5%; Score 139.4; DB 4; Length 2158;
Best Local Similarity 55.1%; Pred. No. 1.5e-30;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

QY 270 CCTGTGAACGTGTTCTCTGGAGTCCCTTTGCTGCTCCCGCGTGGATCCCGCATTT 329
Db 269 CCGTCTCTGCTTCTCTGGAGTCCCTTTGCTGCTCCCGCGTGGATCCCGCATTT 328
QY 330 ACGAACCCGAGCGCTGCATCGCCCTGGAGTAATTGCGAAGCCACTCTCAACCTTAAT 389
Db 329 CTGCCACCGAGCCCAAGACGCTTGTGTAAGGGGTGTAGACGCTCAACCTTTCAGAGT 388
QY 390 TTGTGCTTCAGAACTCAGAGTGGCTGCTTTAGATCAACAATGCTCAAGGTGATTAAC 449
Db 389 GTCTGTACCAATATGTGACACCCCTAACCAGGTTTGAAGGACCGAGATGTGAAC 448
QY 450 CCGAATTCGAGAGTGCAGAACTGCTCTTACCTGAACATCTATGCCCTGCCAGCC 509
Db 449 CCGAACCGTGAAGTGAAGAGTGGCTGCTTACCTGAACATCTATGCCCTGCCAGCC 508
QY 510 GATPACAGCTCCAGCTCCCGCTTGTGAGTGTCCAGAGAGTGGCTTCAAGACTGGC 569
Db 509 CCTACATCCCCA---CCCTGTCTCTGTGTGATCTATGGGGGTGCTTTCAGTGGG 565
QY 570 TCAGCTC-----CATCTTTATGGGTCCGCTGCTGCTGCTTGAAGAGCTGTGCT 623
Db 566 GCTTCTCTTGAAGTGTGATGCGCTTCTTGTGTAACAAGCGAGAGACTGTGCT 625
QY 624 GTGTGCTGCAAGTACCGGCTAGAAATATTTGATTTCTTCAAC---ACATGGGATCAAGAT 680
Db 626 GTGTGCAATGAATACCGGCTAGAAATATTTGATTTCTTCAAC---ACATGGGATCAAGAT 680
QY 681 GCTCCGGGAACTGGGCTTCAAGAGCAGAGTGGCTGTGTCTGTGCTGCTGCTGCTGCTGCT 740
Db 686 GCTCCGGGAACTGGGCTTCAAGAGCAGAGTGGCTGTGTCTGTGCTGCTGCTGCTGCTGCT 745
QY 741 ATCAGATTTCTCGGTGGGAGCCCAAGCTGTGACATCTTTTGGCAGGTCCGGGGAGCC 800
Db 746 GTGGACGCTTCCGGGGTGAACCCGATCAAGTACGCTGTTTGGGAGAACCGGGAGGCC 805
QY 801 ATAAAGTTTCTAATCTTAAATCTGTCTCCATGGCCAAAGCTTATTTCCAAAGCATC 860
Db 806 GCTCTGGGTGAGCACTGTGTGTCTCCGCGCCAGCGGGGCTGTTCACAGAGGCTGTG 865

Oy 861 ATGAGAGTGGGTGGCA 879
Db 866 CTGCAAGCGGTGCCCCA 884

Search completed: June 14, 2005, 23:03:30
Job time : 367.298 secs

This Page Blank (uspto)

QY 1596 TTTGCTCGAACCAGGATCTTAATGGAACGACTTCTCTGTGGCCAGCTTATATCTG 1655
 DB 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuThrProAlaTyrAsnLeu 520
 QY 1656 ACTGAGCAGTACTCTCAGCTGGAATTGAACATGAGCTTGGACAGACTCAAGAACCG 1715
 DB 521 ThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuIysGlnPro 540
 QY 1716 GGGGTGATTTTGAACGACGACCATCCCTGATCTGTGCGCTCCGACATGCTCCAC 1775
 DB 541 ArgValAspPheTyrThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
 QY 1776 AGTCCCTTTCTTCTTAATCTTCTCTCTCTCTCCAGCTTTCTTTTCTTTGTCT 1835
 DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePheCysAla 580
 QY 1836 CCT 1838
 DB 581 Pro 581
 RESULT 2
 ADQ89094 standard; protein; 581 AA.
 ADQ89094;
 ADQ89094;
 21-OCT-2004 (first entry)
 Human urological disorder related protein 53010 SEQ:46.
 urological disorder; uropathic; cytostatic; urinary incontinence;
 benign prostatic hyperplasia; human.
 Homo sapiens.
 WO2004065576-A2.
 05-AUG-2004.
 14-JAN-2004; 2004WO-US000750.
 15-JAN-2003; 2003US-0440318P.
 04-FEB-2003; 2003US-0444783P.
 27-MAR-2003; 2003US-0457901P.
 08-MAY-2003; 2003US-0468775P.
 19-MAY-2003; 2003US-0471614P.
 16-JUN-2003; 2003US-0478742P.
 18-JUL-2003; 2003US-0488529P.
 30-JUL-2003; 2003US-0491156P.
 02-SEP-2003; 2003US-0498594P.
 26-SEP-2003; 2003US-0506332P.
 (MILL-) MILLENNIUM PHARM INC.
 Karicheti V, Silos-Santiago I, Eliasof SD;
 WPI; 2004-562167/54.
 N-PSDB; ADQ89093.
 Use of polypeptides related to urological disorders, e.g. 44390, 54181,
 211 or for identifying a compound capable of treating a urological
 disorder or identifying and treating a subject having a urological
 disorder.
 Claim 1; SEQ ID NO 46; 542pp; English.
 The present invention describes the use of polypeptides related to
 urological disorders for identifying a compound capable of treating a
 urological disorder, identifying a subject having a urological disorder,
 or treating a subject having a urological disorder. Also described: (1) a
 method for identifying a compound capable of treating a urological
 disorder; (2) a method for identifying a subject having a urological

CC disorder; and (3) a method for treating a subject having a urological
 CC disorder. The compound has uropathic and cytostatic activities. The
 CC polypeptides related to urological disorders are useful for identifying a
 CC compound capable of treating a urological disorder, identifying a subject
 CC having a urological disorder, or treating a subject having a urological
 CC disorder. Disorders include urinary incontinence and benign prostatic
 CC hyperplasia. The present sequence represents a human urological disorder
 CC related protein, which is used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 581 AA;
 Alignment Scores:
 Pred. No.: 8,566-291 Length: 581
 Score: 3079.00 Matches: 581
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 79.52% Indels: 0
 DB: Gaps: 0
 US-10-023-515-1 (1-2158) x ADQ89094 (1-581)
 QY 96 ATGCCAAGGAGCTTACTTCAATGCTTCAATGAGTCTTTCTGATTTCCAGCC 155
 DB 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTyrCysPhePheLeuIleLeuGlnPro 20
 QY 156 CTGTTGGGACACAGACAGTGGGGAAGAACTGGGCTTGGCTGAAGGGCCACAGAGAAC 215
 DB 21 LeuLeuGlnYlnAspGlnTyrPheIlyuStrGlyProSerAlaGlnGlyProGlnArgAsn 40
 QY 216 ACCAGGCTGGAGTGAATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGGCTGTG 275
 DB 41 ThrArgLeuGlyTyrIleGlnGlnIlyuSglnValThrValLeuGlySerProAlaProVal 60
 QY 276 AAGTGTCTCTGAGTCCCTTGTGCTGCTCCCGCTGGAGTCCCTGGCATTTACGAAC 335
 DB 61 AsnValPheLeuGlyValIleProPheAlaIleProProLeuGlySerLeuAspPheThrAsn 80
 QY 336 CCGGAGCGTCGATGCGCCGTGGGATTAAGTGGAGAAAGCACTCCATCCATAATTGTGCG 395
 DB 81 ProGlnProAlaSerProIleAspAsnLeuArgIlyuAlaHisSerTyrProAsnLeuGly 100
 QY 396 CTCAGAACTCAGAGTGGCTGCTTGAATCAACATGCTCAAGTGCATTAACCCGAAA 455
 DB 101 LeuGlnAsnSerGlnTyrPheLeuLeuAspGlnHisMetLeuIlyuValHisTyrProIlys 120
 QY 456 TTGGAGTGTCAAGAACTGCTTACTGAACATCTATGCGCTGCCACGCGCATACA 515
 DB 121 PheGlyValSerGlnAspCysLeuTyrLeuAsnIleTyrAlaProAlaHisIleAspThr 140
 QY 516 GGCTCCAAAGTCCCGGTCTTGTGTGCTTCCAGAGAGTGCCTTGAAGTGGCTCAGCC 575
 DB 141 GlySerIlyuLeuProValLeuValTyrPheProGlyGlyAlaPheIlyuThrIlyuSerAla 160
 QY 576 TCCATCTTGAATGAGTCCGCGCTGCTGCTATGAGAGCTGCTGTGTGCTCTCCAG 635
 DB 161 SerIlePheAspGlySerIleAlaLeuAlaIleTyrGlnAspValLeuValIlyuValGln 180
 QY 636 TACCGGCTAGGAATATTTGTTCTTCAACATGAGTCAAGCATGCTCCGGGAACTGG 695
 DB 181 TyrArgLeuGlyIlePheGlyPhePheThrThrTyrAspGlnHisAlaProGlyAsnThr 200
 QY 696 GCGTTCAAGACACAGTGGCTGCTTGTGCTGGCTCCGAGGAAGAAATATGAGTCTTGGT 755
 DB 201 AlaPheIlyuAspGlnValAlaIleLeuSerTyrValGlnIlyuAsnIleGlnPhePheGly 220
 QY 756 GGGGACCCAGCTGTGACCATCTTTGGCAGTCCGCGGAGGCATAGTGTCTTAGT 815
 DB 221 GlyAspProSerSerValThrIlePheGlyIleuSerAlaGlyAlaIleSerValSerSer 240
 QY 816 CTATATCTGTCTCCATGAGCCAAAGGCTTATTCACAAAGCATATGAGAGTGGGCTG 875
 DB 241 LeuIleLeuSerProMetAlaIlyuGlyLeuPheHisIlyuAlaIleMetGlySerGlyVal 260

```

QY      876 GCCATATCCCTTACCTGGAGGCCCATGATTAAGAGAGTGAAGACCTCGAGGTGGTT      935
      261 AAlaIleIleProTyrLeuGluuAlaHisAspTyrGluuYsSerGluuAspLeuGluValVal      280
QY      936 GCACATTTTGTGTGTAACAATGCGTCAGACTGTGAGGCCCTGTGAGGCGCTTGAGGACA      995
      281 AAlaHisPheCysGluYAsnAsnAlaSerAspSerGluuAlaLeuLeuHisCysLeuArgThr      300
QY      996 AAACCCCTCAAGAGAGCTGTGACCTCGACCCGAAAGAACAAAGCTTTTCTACTCGAGTGGTT      1055
      301 YsPProSerLysGluLeuLeuLeuThrLeuSerGlnuYsThrLysSerPheThrArgValVal      320
QY      1056 GATGTGCTTTTCTTCTTAATAGCCCTGATGATCTATTGTGTGCAAGAAAGCTTTTAAGCA      1115
      321 AArgGlyAlaPhePheProAsnGluuProLeuAsnProLeuLeuSerGlnuYsAlaPheLysAla      340
QY      1116 ATTCCTTCATCATCATGGAGTCAATTAACAAGAGTGTGGCTTCTGTGCTGCTATGAGAG      1175
      341 ILeProSerIleIleGlyValAlaAsnAsnHisGluCysGluPheLeuLeuProMetLysGlu      360
QY      1176 GCTCTGAGATCTCAGTGTGCTCCAAAGAGTCCCTGCGCCCTGCTGATTAACAAATC      1235
      361 AlaProGluIleLeuSerGlySerHisnLysSerLeuAlaLeuHisLeuIleGlnAsnIle      380
QY      1236 CTGACATCCCGCTCCAGTATTGTCACCTTGTGGCTAATGATCTTCATGACAAAGAC      1295
      381 LeuHisIleProProGlnuYrLeuHisLeuValAlaAsnGluYrPheHisAspLysHis      400
QY      1296 TCCCTGACATGAATCCGAGACAGTCTTGGACTGTGAGATGTGCTTCTTGTGGCTC      1355
      401 SerLeuThrGluIleArgAspSerLeuLeuAspLeuLeuGluYAspAlaPhePheValVal      420
QY      1356 CTGACATGATCACAGTCTGATATCACAGAGATGCTGTGTCACCTGTCTACTTCTATGAG      1415
      421 ProAlaLeuIleThrAlaArgTyrHisArgAspAlaGlyAlaProValYrPheTyrGlu      440
QY      1416 TTTCGGACCGGCTCAGTCTTTGAAAGACAGAACCGGCTTTTGTCAAGCCGACAC      1475
      441 PheArgHisArgProGlnuYsPheGluAspThrLysProAlaPheValLysAlaAspHis      460
QY      1476 GCTGATGAGATCGCGCTTTGTGTGGTGTGCTTCTGTAAGGGGAGACATGTGTAATGTT      1535
      461 AlaAspGluValAlaArgPheValAPheGlyGlyAlaPheLeuLysGluYAspIleValMetPhe      480
QY      1536 GAAGAGCCACGGAGAGAGAGAACTTACAGACCGGAAAGATGATGAATATCTGGCTTACC      1595
      481 GluGlyAlaThrGluGluGluGluLeuLeuSerArgLysMetMetLysTyrTrpAlaThr      500
QY      1596 TTTCCTGACACCGGATCTTATGGAGACGACTGTCTGTGGCCAGCTTATTAATCTG      1655
      501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuIleTrpAlaTyrAsnLeu      520
QY      1656 ACTGAGCAGTACTCAGCTGAGTGAACATGAGCCTTGGACAGAGACGTAAGAACCG      1715
      521 ThrGluGlnuYrLeuGlnLeuAspLeuAsnMetSerLeuGlnArgLysLeuGluPro      540
QY      1716 CCGGTGATTTTGGACGACGACATCCCGCTGATCTGTGCTGCTCGAATAGTCTCAC      1775
      541 ArgValaAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis      560
QY      1776 AGTCTCTTTTCTTCTTAACTTCTCTCTCTCTCCAGCCTTCTTTTCTTGTGTGT      1835
      561 SerProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePhePheCysAla      580
QY      1836 CCT 1838
      581 Pro 581

```

RESULT 3
 ABP61004
 ID ABP61004 standard; protein; 575 AA.
 .xx

```

AC      ABP61004;
XX
XX      10-SEP-2002 (first entry)
DT
XX
DE      Novel human protein. SEQ ID 91.
KW
KW      Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW      nootropic; neuroprotective; immunosuppressive; haemostatic;
KW      antiinflammatory; cardiant; antitumor; vitruide; antithyroid;
KW      cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW      wound healing disorders; atherosclerosis; Parkinson's disease;
KW      Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW      inflammation; neoplastic diseases; nervous system disorder;
KW      hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW      developmental abnormality; gastrointestinal ulceration; neuropathy;
KW      haematological disease; metabolic disease; sperm dysfunction;
KW      thyroid disorder; hypothyroidism; brain damage; colitis;
KW      cone photo-transduction deficiency; neurological disease; stroke;
KW      anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW      trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW      growth abnormality; precocious puberty.
OS
XX      Homo sapiens.
XX
XX      WO200250105-A1.
PN
XX
XX      27-JUN-2002.
PD
XX
XX      17-DEC-2001; 2001WO-US049232.
PF
XX
XX      19-DEC-2000; 2000US-0256710P.
PR
XX      20-DEC-2000; 2000US-0257048P.
PR      09-JAN-2001; 2001US-0260482P.
PR      30-JAN-2001; 2001US-0264922P.
PR      06-FEB-2001; 2001US-0266797P.
PR      19-MAR-2001; 2001US-0276988P.
PR      04-APR-2001; 2001US-0281535P.
PR      08-MAY-2001; 2001US-0289622P.
XX
PA      (SMIK ) SMITHKLINE BEECHAM CORP.
PA      (SMIK ) SMITHKLINE BEECHAM PLC.
PA      (GLAX ) GLAXO GROUP LTD.
XX
PI      Agrawal P, Birkeland M, Cogswell JP, Kahnrick KF, Lai Y;
PI      Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX      WPI; 2002-508764/54.
DR      N-PSDB; AB086169.
XX
XX      Secreted proteins and polynucleotides useful as vaccines for preventing
XX      or treating various diseases e.g. cancer, wounds, atherosclerosis,
XX      Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
XX      Claim 1(a); Page 312-313; 335pp; English.
XX
XX      The invention relates to an isolated polypeptide with signal sequences
XX      which allow it to be secreted extracellularly or membrane associated. The
XX      activity of polypeptides of the invention may be described as,
XX      cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, nootropic,
XX      neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
XX      cardiant, antitumor, vitruide, antithyroid, cerebroprotective, anorectic,
XX      and metabolic. Polypeptides and polynucleotides of the invention are
XX      useful in the treatment, or as a vaccine in the prevention of, cancer,
XX      wound healing disorders, infection, atherosclerosis, Parkinson's disease
XX      and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
XX      inflammation, neoplastic diseases, nervous system related disorders and
XX      cardiovascular disorders, pancreatitis, respiratory disorder,
XX      hyperproliferation, systemic autoimmune disease, hyper-immunity,
XX      developmental abnormality, gastrointestinal ulceration, neuropathy,
XX      haematological diseases, metabolic diseases, sperm dysfunction, thyroid
XX      disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
XX      transduction deficiency, neurological diseases, stroke, anglogenesis.

```


proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;
 aetmia; neurological disorder; Alzheimer's disease; Huntington's disease;
 dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;
 drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;
 renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;
 anoxemia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;
 goltre; gastrointestinal disorder; gene therapy; viricide; anticoagulant;
 anticonvulsant; nootropic; enzyme; DME-10.

Homo sapiens.

Key Location/Qualifiers
 Peptide 1..50
 /label= Signal_peptide
 Domain 32..56
 /note= "Transmembrane domain"
 Protein 51..642
 /note= "Mature human DME-10"
 Domain 113..135
 /note= "Transmembrane domain"
 Domain 204..220
 /note= "Transmembrane domain"
 Domain 234..250
 /note= "Transmembrane domain"
 Domain 287..314
 /note= "Transmembrane domain"
 Domain 463..491
 /note= "Transmembrane domain"

MO200246426-A2.

13-JUN-2002.

04-DEC-2001: 2001MO-US047429.

08-DEC-2000: 2000US-0254308P.

15-DEC-2000: 2000US-0256189P.

21-DEC-2000: 2000US-0257713P.

19-JUN-2001: 2001US-0262706P.

02-FEB-2001: 2001US-0266020P.

(INCY-) INCYTE GENOMICS INC.

Sanjwal MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ,
 Lee EA, Ding L, Hafalla AJA, Tang YT, Yue H, Trilouley CM, Lu DM,
 Lal PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AR, Lu Y,
 Ison CH;

WPI: 2002-519668/55.
 N-PSDB: AAD40574.

Novel human drug metabolizing polypeptide, useful in diagnosis,
 prevention or treatment of autoimmune/inflammatory, cell proliferative,
 neurological, developmental, endocrine, metabolic and gastrointestinal
 disorders.

Claim 65; Page 155-156; 169pp; English.

The invention relates to an isolated human drug metabolising enzyme (DME)
 and its nucleotide. DME is useful for diagnosing, treating or preventing
 disorders associated with aberrant expression of DME, where the disorders
 are selected from autoimmune/inflammatory disorder such as acquired
 immunodeficiency syndrome (AIDS), asthma, arteriosclerosis, psoriasis,
 uveitis; a cell proliferative disorder such as arteriosclerosis,
 cirrhosis, hepatitis, and cancer; a neurological disorder such as
 Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;
 anemias; an endocrine disorder such as adenoma, thrombosis and infections
 ; an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic
 disorder such as cystic fibrosis, diabetes and goltre; a gastrointestinal
 disorder such as anorexia, peptic ulcer; and liver disorders. DME is
 useful in a number of drug screening techniques and to analyse the
 proteome of a tissue or cell type. The invention is useful for creating

CC knock-in humanised animals or transgenic animals to model human diseases,
 CC in somatic or germline gene therapy, to generate a transcript image of a
 tissue or cell type, for detecting differences in the chromosomal
 CC location due to translocation, inversion, etc. among normal, carrier or
 affected individuals, and as hybridisation probes for mapping naturally
 CC occurring genomic sequences. The present sequence is human DME-10
 XX
 SQ Sequence 642 AA;

Alignment Scores:

Pred. No.: 2 74e-273 Length: 642
 Score: 2800.00 Matches: 548
 Percent Similarity: 99.82% Conservative: 2
 Best Local Similarity: 99.46% Mismatches: 1
 Query Match: 74.90% Indels: 0
 DB: 5 Gaps: 0

US-10-023-515-1 (1-2158) x AAE25025 (1-642)

QY 186 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTACAGGCGACAA 245
 Db 92 GTPProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyValGln 111
 QY 246 GTCACTGTGGTGGGAAAGCCCTGTGCTGTGAACGTGTTCTCGAGTCCCTTGTGCTG 305
 Db 112 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 131
 QY 306 CCCCCGTGGGATCCCTGCAGATTACGACCCGAGCTGCATCGCCCTGGAGTAACCTTG 365
 Db 132 ProProLeuGlySerLeuAlaGlnPheThrAsnProGlnProAlaSerProTrpAspAsnLeu 151
 QY 366 CGAAGACCAACCTTCCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTCTTAGAT 425
 Db 152 ArgGlnAlaThrSerGlyProAlaLeuGlyLeuGlnAsnSerGlyTrpLeuLeuAsp 171
 QY 426 CAACACATGCTCAAGGTGATTAACCCGAATTGAGAGTGCAGAGACAGCTGCTCACTG 485
 Db 172 GlnHisMetLeuGlyValHisGlyTrpGlyPheGlyValSerGlnAspGlyLeuValTrpPhe 191
 QY 486 AACATCTATGCGCTGCGCCACGCGGATACAGGCTCCAGCTCCCGCTGTGGTGTGTC 545
 Db 192 AsnIleTrpAlaProAlaHisAlaAspTrpGlySerGlyLeuProValLeuValTrpPhe 211
 QY 546 CCAGAGAGTCTCTCAAGAGCTGCTCAGGCTCATCTTTGATGGTCCGCTGCTGCTGCC 605
 Db 212 ProGlyGlyAlaPheGlyThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 231
 QY 606 TATGAGACGCTGCTGCTGTGCTGCTGCATACCGGCTAGAAATATTGGTTCTTCACC 665
 Db 232 TyrGlnAspValLeuValValValGlnTrpArgLeuGlyIlePheGlyPhePheThr 251
 QY 666 ACATGGAGTCAGATGCTCCGGGAACTGGGCTTCAGAGACAGAGTGGCTGCTGCC 725
 Db 252 ThrTrpAspGlnHisIleAlaProGlyAsnTrpAlaPheGlyAspGlnValAlaAlaLeuSer 271
 QY 726 TGGGTCCAGAGAACATCGAGTTCTTCGGTGGGAGCCAGGCTGTGATCCATCTTGGC 785
 Db 272 TrpValGlnIleAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGly 291
 QY 786 GAGTCCGCGGAGCCATTAAGTGTTCATGCTTATTAAGTGTTCATGAGGCGCAAGGCTTA 845
 Db 292 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPrometAlaGlyGlyLeu 311
 QY 846 TTCACAAAGCCATCATGAGAGAGTGGGAGCCATCATCCCTTACCTGAGGCGCCATCAT 905
 Db 312 PheHisIleValAlaIleMetGlnSerGlyAlaIleIleIleProGlyLeuGlnAlaHisAsp 331
 QY 906 TATGAGAGAGTGAAGACCTCGAGGTGTTGCATTTCTGTGTAACAATGCTGTCAGAC 965
 Db 332 TyrGlnIleValSerGlnAspLeuGlnValValAlaHisPheCysGlyAsnAlaAspSerAsp 351
 QY 966 TCTGAGGCTCTGAGAGTGTGCTGAGACAAACCTTCAGAGAGTGTGCTGACCTGACG 1025


```

Db      352 SerGluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSer 371
QY      1026 CAGAAAACAAGCTTTTCACTGAGAGTGTTGATGAGCTTTTCTTCTAATAGAGCTCTA 1085
Db      372 GlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeu 391
QY      1086 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATGATCCGAGTCAATACAC 1145
Db      392 AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyAlaAsnAsnHis 411
QY      1146 GAGTGTGCTTCTGCTGCTGCTTGAAGAGGCTCTGAGATCCTCAAGTGGCTTCAACAG 1205
Db      412 GlnCysGlyPheLeuLeuProMetLysGluAlaProGluValLeuSerLysSerAsnLys 431
QY      1206 TCCCTTGGCTCCATCTGATAGAAACATCCGACATCCCGCTCAGATTGTGACCTT 1265
Db      432 SerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIleuHisLeu 451
QY      1266 GTGGCTAATGAACTACTCCATGACAAAGCACTCCCTGACTGAATCCGAGACAGTCTTCTG 1325
Db      452 ValAlaAsnGluTrpPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeu 471
QY      1326 GACTTGTCTTGAGATGTGTCTTGTGTGCTCTGCACTGATCAAGCTCGATTATCAGAGA 1385
Db      472 AspLeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTrHisArg 491
QY      1386 GATGTGTGGGACCTGTCTACTTCTATGAGTTTCCGACCGGCTTCAGTCTTGAAGAC 1445
Db      492 AspAlaGlyAlaProValTrpPheTrpGluPheAspGlnHisArgProGlnCysPheGluAsp 511
QY      1446 ACGAAGCCGGCTTTTGTCAAAAGCCGACAGCTGATGAGTCCGCTTGTGTGGTGGT 1505
Db      512 ThrLysProAlaPheValLysAlaAspHisAlaAspGluValArgPheValPheGlyGly 531
QY      1506 GCTTCTCTGAGAGGGGACATTGTATGTTGAAGAGCCAGAGAGGAGAGAAATTACTG 1565
Db      532 AlaPheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuLeu 551
QY      1566 AGCCGGAAGATGATGAAATATCGGGTACTTGTCTCGAAGCCGGAATCTTATGGGAAC 1625
Db      552 SerArgLysMetMetLysTrpAlaThrPheAlaArgThrGlyAsnProAsnGlnLysAsn 571
QY      1626 GACTGTCTCTGTGGGACGCTTATATGACTGAGACGACTTCCAGCTGAGCTTGAAC 1685
Db      572 AspLeuPheLeuTrpProAlaTrpAsnLeuThrGluGlnTrpLeuGlnLeuAspLeuAsn 591
QY      1686 ATGAGCTCTGGAACAGACTCAAGAACCCGCGGTGGATTATTGGAACAGACCATCCCC 1745
Db      592 MetSerLeuGlyGlnArgLeuLysGluProArgValGluPheTrpHisSerThrIlePro 611
QY      1746 CTGATCTCTGTGCTCTCGACATGCTCCAGAGTCTTCTTCTTAACTTCTCTCT 1805
Db      612 LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer 631
QY      1806 CTCTCCAGCGCTTCTTCTTCTTGTGCTCC 1838
Db      632 LeuLeuGlnProPhePhePheCysAlaPro 642

```

```

RESULT 5
ABM84114
ID      ABM84114 standard; protein; 642 AA.
XX      ABM84114;
AC      18-NOV-2004 (first entry)
DT      Human diagnostic and therapeutic pprotein SEQ ID NO:4363.
XX      Human diagnostic and therapeutic pprotein SEQ ID NO:4363.
DE      gene therapy; human diagnostic and therapeutic polynucleotide; dtltp.
XX      Homo sapiens.
OS      Homo sapiens.
XX      WO2004023973-A2.
PN

```

```

XX      25-MAR-2004.
PD      12-SEP-2003; 2003WO-US028227.
XX      12-SEP-2002; 2002US-0410259P.
XX      12-SEP-2002; 2002US-0410259P.
PR      12-SEP-2002; 2002US-0410260P.
XX      (INCYTE CORP.
PA      Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
XX      Hachtshorne TA, Suchorolski MT, Altus CM, Plets SJ, Elder LV,
PI      Mooney EM, Deleage AM, Panesar IS, Banyalle SC, Reddy TP,
PI      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH,
PI      Perla CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LI,
PI      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Velt UA, Kirton BS,
PI      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D,
PI      Fatuty S, Shi X, Suarez CJ,
XX      MPI; 2004-329368/30.
DR      N-PSDB; ACN42766.
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT      in diagnosing a condition, disease or disorder associated with human
PT      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT      in gene mapping.
XX      Claim 27, Page: 190pp; English.
PS      The invention relates to novel diagnostic and therapeutic polynucleotides
XX      selected from one of the 2722 sequences defined in the specification. A
CC      polynucleotide of the invention may have a use in gene therapy. The human
CC      diagnostic and therapeutic polynucleotides (dtltp) or polypeptides may be
CC      used to diagnose a particular condition, disease or disorder associated
CC      with human molecules, e.g. cell proliferative disorders,
CC      autoimmune/inflammatory disorders, developmental disorder, endocrine
CC      disorder, neurological disorders, gastrointestinal disorders, or
CC      infections caused by virus, bacteria, fungi or parasite. The dtltp
CC      molecules may also be used in genetic mapping, in identifying individuals
CC      from minute biological samples, in detecting single nucleotide
CC      polymorphisms, as molecular weight markers, and for somatic or germline
CC      gene therapy. The present sequence data for this patent is not represented in
CC      the printed specification, but was obtained in electronic format directly
CC      from WFO at www.wipo.int/pct/en/sequences/listing.htm
XX      Sequence 642 AA;
SQ

```

```

Alignment Scores:
Pred. No.: 2,74e-273 Length: 642
Score: 2800.00 Matches: 549
Percent Similarity: 99.82% Conservative: 1
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 74.90% Indels: 0
DB: 8 Gaps: 0

```

```

US-10-023-515-1 (1-2158) x ABM84114 (1-642)
QY      186 GGGCTTCTGTGAAGGGGCAAGAGAAACACAGGCTGGATGATTCAGGGCAAGAA 245
Db      92 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnLysGln 111
QY      246 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTCTCTGGAGTCCCTTGTGCTGT 305
Db      112 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 131
QY      306 CCCCGCTGGATCCCTGCGATTTTACGAACCCGACAGCTGATCCGCTGGATTAATTG 365
Db      132 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeu 151
QY      366 CGAAGACCACTCTTACCTTAATTGTGCTCCAGAACTCAGAGTGGCTCTTATGAT 425
Db      152 ArgGluAlaThrSerTrpAsnLeuCysLeuGlnAsnSerGluTrpLeuLeuLeuAsp 171

```

QY	426	CACACATGCTCAAGTGCATATCCGAAATTCGGAGTGCAGAAAGATGCTGCTACCTG	488
Db	172	GlnHsMestLeuLyValHnIstYrProLysPheGlyValSerGlnAAsrCySleuTyLeu	191
QY	486	AACATCTATGGCGCCGCCCAAGCGGATGCAGGCTCCAAAGTCCCGCGCTTGGTGGTTC	545
Db	192	AenIleTyxAlaProLanIsAlaAsrPhtGlySerTyLeuProValLeuValTyrPhe	211
QY	546	CCAGAGGTGCTTCAGAGCTGGTGCAGCTCCCAATCTTTGATGGGTGCGCTGGCTGCC	605
Db	212	ProGlyGlyAlaPheLyTyHcGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla	231
QY	606	TATAGAGACGGTGGCTGTTGGTGGTGCATGACCGGCTAGCAATATTGGTTCTTCACC	665
Db	232	TyrGlnAsrValLeuValAlaValaGlnTyxArgLeuGlyIlePheGlyPhe***Ser	251
QY	666	ACATGGGATGAGCATGCTCCGCGGGAATGGAGCCCTCAAGGACGAGTGGTCTGTGTC	725
Db	252	ThrTyrAspGlnHnIsAlaProGlyAsnTyrAlaPheLysAspGlnValAlaAlaLeuSer	271
QY	726	TGGGTCCAGAAAGACATGAGTTCTTCGATGGGAGCCCGACGTCTGTGACATCTTTGAC	785
Db	272	TyrValGlnLysAsnIleGlyPhePheGlyGlyAspProSerSerAlaThrIlePheGly	291
QY	786	GAGTCCGCGGAGACCATTAAGTGTTCATGCTTATCTGTCTCCATGCGCAAGGCTTA	845
Db	292	GlnSerHnIleGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu	311
QY	846	TTCCAAAGGCCATCGAGAGAGTGGGTGGCATCATCTTGTGGTAAACATGCGTCAGAC	905
Db	312	PheHnIleValAlaIleMetGlnSerGlyValAlaIleIleProTyLeuGlnAlaHnIsAsp	331
QY	906	TATGAGAAAGTGAAGACCTGCAGGGTGGTGCACATTTGTGGTAAACATGCGTCAGAC	965
Db	332	TyrGlnLysSerGlnAsrLeuGlnValValaLanIsPheCyGlyAsnAsnAlaSerAsp	351
QY	966	TGTGAGCCCTGCTGAGTGGTGCAGAGCAAAACCTCCAGAGACCTGTACCTCAGC	1025
Db	352	SerGlnAlaLeuLeuHnIleGlyCySleuHnArgThrLysProSerLysGlnLeuHnIleHnSer	371
QY	1026	CAGAAAACAAAGTCTTTCACTGAGTGGTGAATGGTCTTTCTTCTAATAGACCTTA	1085
Db	372	GlnLysTyrLysSerPheThrArgValAlaAspGlyAlaPhePheProAsnGlnProLeu	391
QY	1086	GATCATATGTCTCAGAAAGCATTTTAAACAAATTCCTTCATCATCGGAGTCAATTAACAC	1145
Db	392	AspLeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValaAsnAsnHnIs	411
QY	1146	GAGTGGCTTCTGCTGCTGATGAGAGAGCTCTGAGATCTCAGTGGCTCAACAG	1205
Db	412	GlnCySbeGlyPheLeuLeuProMetLysGlnAlaProGlnIleLeuSerGlySerHnLys	431
QY	1206	TCCCTGCGCCCTCAATCGATCAAAACATCTGCAATCCGCGCTCAAGTATTTGACSTT	1265
Db	432	SerLeuAlaLeuHnIsleuIleGlnAsnIleLeuHnValleProProGlnTyLeuHnIsleu	451
QY	1266	GTGGCTAATGATATCTTCATGACAAAGACACTCCCTGATAAAATCCAGACAGTCTTGC	1325
Db	452	ValAlaAsnGlnTyPheHnIsAsrLysHnIsSerLeuHnIleArgAspSerLeuLeu	471
QY	1326	GACTTGGTTGAGATGTTCTTTGTGGTCCCTGCACATGCACAGCTGCATATCAAGA	1385
Db	472	AspLeuLeuLysAsrValaPhePheValaProAlaLeuIleThrAlaArgTyHnIsArg	491
QY	1386	GATGCTGGTGCACCTGCTCATCTTATGAGTTTCGGACCGGCGCTCAAGCTTTGAAGAC	1445
Db	492	AspAlaGlyAlaProValTyHnIleTyxGlnPheArgHnIsArgProGlnCyPheGlyAsp	511
QY	1446	ACGAAAGCCGGCTTTGTCAAAAGCCACACAGCTGATGAATCGCTTGTGTCGGTGGT	1505
Db	512	ThrLysProAlaPheValLysAlaAspHnIsAlaAspGlnValaArgPheValaPheGlyGly	531

QY	1506	GCCTTCCTGAAGGGGACATGTGTTATGTTGAGAGGACGACGAGAGAGAGATTACTG	1565
Db	532	AAlapheuleuysciIaepileValmetPheglungIylAlathGlungIulbIeuleu	551
QY	1566	AGCCGGAAGATGANTGAATATCTGGGCTACTTTTGCTCGAACCGGGAATCTTAATGGGAAC	1625
Db	552	SerArglysmecIeysIyrtIrlPalatIrlPhealIaAgYthIrgIYAsnPrOAengIYAsn	571
QY	1626	GACCTGTCTCTGGCCAGCTTATATGATGAGCTGAGCACTCCACCGCTGAGACTTGAAC	1685
Db	572	AspIeuSerIeuIrlProAlaIrlZhenIeuIrlGluIrlYIrlGluIrlLeuAapIeAAsn	591
QY	1686	ATGAGCCTCGGAGAGAGACTCAAGAGACCGCGGCTGATTTTGGACGAGACCATCCCC	1745
Db	592	MetSerIeuIylGluIrlArgIleuysgluIrlProArgValaIaPheIrlPrlSerIrlIlePro	611
QY	1746	CTGATCCGTGTCTGCCCTCCAGATGGCTCCACAGTCTCTTTTTCCTTAATCTTCTCTCTC	1805
Db	612	LeuIleIeuSerIrlAserIrlPheIeuIrlSerIrlProIeuSerIeuIrlPheIeuSer	631
QY	1806	CTCCTCAGCCTTTCTTTTCTTTTGTGCTCCT	1838
Db	632	LeuIeuGluIrlProPhePhePheCysIaIrlPro	642
RESULT 6			
ID	ABP61005	standard; protein; 581 AA.	
XX	AC	ABP61005;	
XX	DT	10-SEP-2002 (first entry)	
XX	DE	Novel human protein. SEQ ID 92.	
XX	KW	Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;	
KW	nootropic; neuroprotective; immunosuppressive; haemostatic;		
KW	antiinflammatory; cardiac; antitumor; vitruide; antihypoid;		
KW	cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;		
KW	wound healing disorders; atherosclerosis; Parkinson's disease;		
KW	Alzheimer's disease; autoimmune disorders; haematopoietic disorder;		
KW	inflammation; neoplastic disease; nervous system disorder;		
KW	cardiovascular disorders; pancreatitis; respiratory disorder;		
KW	hyperproliferation; systemic autoimmune disease; hyper-immunity;		
KW	developmental abnormality; gastrointestinal ulceration; neuropathy;		
KW	haematological disease; metabolic disease; apem dysfunction;		
KW	thyroid disorder; hypothyroidism; brain damage; colitis;		
KW	cone photo. transduction deficiency; neurological disease; stroke;		
KW	anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;		
KW	trachea; thymus; lymph node; muscular system; obesity; anorexia;		
KW	growth abnormality; precocious puberty.		
XX	OS	Homo sapiens.	
XX	PN	WO200250105-A1.	
XX	PD	27-JUN-2002.	
XX	PF	17-DEC-2001; 2001WO-US049232.	
XX	PR	19-DEC-2000; 2000US-0256710P.	
PR	20-DEC-2000; 2000US-0257048P.		
PR	09-JAN-2001; 2001US-0260482P.		
PR	30-JAN-2001; 2001US-0260482P.		
PR	06-FEB-2001; 2001US-0266979P.		
PR	19-MAR-2001; 2001US-0276988P.		
PR	04-APR-2001; 2001US-0281535P.		
PR	08-MAY-2001; 2001US-0286922P.		
XX	PA	(SMIK) SMITHKLINE BEECHAM CORP.	
PA	(SMIK) SMITHKLINE BEECHAM CORP.		
PA	(GLAX) GLAXO GROUP LTD.		
XX	Agarwal P, Birkeland M, Cogswell JB, Kahnrick KE, Tai V,		


```

OY 1269 GCTATGAAATCTTCATGACACAGACTCCCTGACTGAAATCCGAGACAGTCTTGAC 1328
DB 386 A|aa|eng|ur|y|Phe|hi|b|asp|y|shi|s|ser|Leu|thr|g|lu|l|e|arg|b|ser|Leu|asp 405
OY 1339 TTGCTTGAGATGATGTTCTTTGAGTCCCTGACATGATCAGAGCTCGATATGACAGAGAT 1388
DB 406 Leu|leu|g|ly|A|asp|Val|Phe|b|he|Val|Val|Pro|al|e|u|l|e|Thr|Al|a|arg|y|hi|b|arg|--- 424
OY 1389 GCTGTGACCTGTCTACTTCTATAGATTTCGGCACCGGCTCAAGTCTTGAACACAG 1448
DB 424 ----- 424
OY 1449 AAGCGGCTTTTGTCAAGCGACACCGCTGATGAGTCCGCTTGTGTTCCGTGTGTC 1508
DB 424 ----- 424
OY 1509 TTCCTGAAGGGGACATTGTTATGTTGAGAGGCCACGAGAGAGAGAGATTACTGAGC 1568
DB 425 -----glu|g|ly|Ala|Thr|glu|glu|glu|yls|Leu|Leu|ser 435
OY 1569 CGAAGATGATGAAATTA|CTGGGCTACCTTGTCTGACACCGGGAATCTTATGGGAACGAC 1628
DB 436 Arg|lys|Met|Met|yls|Tyr|Tyr|Ala|Thr|Phe|Ala|Arg|Thr|gl|y|Asn|Pro|b|n|gl|y|Asn|b|asp 455
OY 1629 CTGTCTGTGGCCGACTTATATCTGACCTGACAGTACCTCCAGCTGAGCTTGAACATG 1688
DB 456 Leu|ser|Leu|Tyr|Pro|Ala|Tyr|Asn|Leu|Thr|glu|g|ly|Tyr|Leu|g|ln|Leu|asp|Leu|b|n|Met 475
OY 1689 AGCTTCGACAGACAGACTCAAGACCGCGGTGATGATTTTGGACAGACACATCCCCCTG 1748
DB 476 Ser|leu|g|ly|g|ln|Arg|Leu|yls|glu|Tyr|Pro|Arg|Val|Asn|Phe|Thr|Thr|Ser|Thr|lle|Pro|leu 495
OY 1749 ATCTGTCTGCTCTCGACAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTCTC 1808
DB 496 lle|Leu|Ser|Ala|Ser|Asp|Met|Leu|His|Ser|Pro|Leu|Ser|Ser|Leu|Thr|Phe|Leu|Ser|Leu 515
OY 1809 CTCGAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1838
DB 516 Leu|g|ln|Pro|Phe|Phe|Phe|b|Cys|Ala|Pro 525

RESULT 8
ABUS4639
ID ABUS4639 standard; protein; 581 AA.
XX
AC ABUS4639;
XX
DT 03-JUN-2003 (first entry)
XX
DE Human NOVX polypeptide #98.
XX
KW Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atriocentricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX
OS Homo sapiens.
XX
PM WO200281498-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002MO-US010780.
XX
PR 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.

```

```

PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283513P.
PR 13-APR-2001; 2001US-02837210P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332792P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.

(CUBA-) CUBAGEN CORP.
XX
FA Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek RA;
XX Paturnajan M, Liu X, Gusev VY, Li L, Vernet CM, Zethusen BD;
XX Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
XX Padigaru M, Shinkets RA, Gangoli EA, Taupier RJ, Caeman SJ, Ji W;
XX Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
XX MacDougall JR, Rothenberg MB, Mazur A, Millet I, Peyman JA;
XX Ellerman K;
XX MPI: 2003-046858/04.
XX N-PSDB: ABX72267.
XX
PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 303; 666p; English.
XX
XX The invention relates to human polypeptides, termed NOVX, and the
XX polynucleotides encoding them. The polypeptides and polynucleotides are
XX useful for diagnosing disease, and screening for potential therapeutic
XX agents. The sequences are useful for treating metabolic disorders,
XX cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
XX stenosis, atrial septal defect (ASD), atriocentricular canal defect,
XX ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
XX septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
XX atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
XX disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
XX haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
XX and cancer. Sequences ABUS4542-ABUS4647 represent human NOVX polypeptides
XX of the invention
XX
SQ Sequence 581 AA;
XX
Alignment Scores:
Pred. No.: 8,966-244 Length: 581
Score: 2597.50 Matches: 497
Percent Similarity: 95.07% Conservative: 4
Best Local Similarity: 94.31% Mismatches: 17

```

Query Match: 67.08% Indels: 9
DB: 6 Gaps: 3
US-10-023-515-1 (1-2158) x ABUS4639 (1-581)

```
QY 186 GGGGCTTGTGTGAAGGGCCACAGAGAAACACAGAGTGTGGATTGATTCAGGGCAAGCA 245
DB 27 G1YProSerAlaGluGlyProGlnAArgSerThrArgLeuGlyTrpIleGlnGlyLysGln 46
QY 246 GTCACTGTGTGGGAAGCCCTGTGCTGTGAAGCTGTCTCGGAGTCCCTTTGCTGTCT 305
DB 47 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla 66
QY 306 CCCCCTGGGATCCCTGCGATTACGAACCCGAGCCGTCATGCGCTGGGATTAATTG 365
DB 67 ProProLeuGlySerLeuAArgPheThrAsnProGlnProAlaSerProTrpAsnLeu 86
QY 366 CGAAGACCCACCTCTTACCTTAATTGTGCTCCAGAACTCAGAGTGGCTGCTTAAAT 425
DB 87 ArgGluAlaThrSerTyProAsnLeuCysLeuGlnAsnSerGluTrpLeuLeuAsp 106
QY 426 CAACACATGCTCAAGGTCATTACCGAAATTCGAGATGCAGAAAGCTGCTTACCTG 485
DB 107 GlnHisMetLeuLysValHisTyProLysPheGlyValSerGluAspCysLeuTyLeu 126
QY 486 AACATCTATGCGCTGCGCCACAGCCGATACAGGCTCCAGGCTCCCGCTTGTGTGCTTC 545
DB 127 AsnIleTyAlaProAlaHisAlaAspThrGlySerTyLeuProValLeuValTrpPhe 146
QY 546 CCAAGAGGTGCTTCAAGACTGGCTCGATCTTGAATGGTCCGCTGTGCTGCTCC 605
DB 147 ProGlnGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla 166
QY 606 TATGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
DB 167 TyGluAspValLeuValValValGlnTyArgLeuGlyTyrPheGlyPhePheThr 186
QY 666 AACTGGGATCAGATCTCTCCGGGAACTGGGCTTCAAGACCCAGGTGGCTGTGTCTCC 725
DB 187 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSer 206
QY 726 TGGGTTCAGAGAAACATCGAGTTCTTCCGTTGGGGACCCGAGCTGTGTACATCTTTGGC 785
DB 207 TrpValGlnLysAsnIleGluPhePheGlyGlyAspProSerSerValThrIlePheGly 226
QY 786 GAGTCCCGGGAGCCATTAAGTGTCTTAACTGTATCTGTCCATGAGCCAAAGCTTAA 845
DB 227 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeu 246
QY 846 TTCACAAAGCCATCATGAGAGTGGGGTGGGACATCCCTTACCTGAGAGCCCATGAT 905
DB 247 PheHisLysValAlaIleMetGluSerGlyAlaAlaIleIleProTyLeuGlnAlaHisAsp 266
QY 906 TATGAGAAAGTGAAGACCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 965
DB 267 TyGluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnHisSerAsp 286
QY 966 TCTGAGGCGCTGTGAGGTGCTGAGGACAAACCTTCAGAGGCTGCTGACCTTACG 1025
DB 287 SerGluAlaLeuLeuAspCysLeuAspArgThrLysProSerLysGluLeuLeuThrLeuSer 306
QY 1026 CAGAAAACAAAGTCTTTCATCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1085
DB 307 GlnLysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeu 326
QY 1086 GATCATTTGTCTCAAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1145
DB 327 AsnLeuLeuSerGlnLysValAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHis 346
QY 1146 GAGTGTGCTTCTGCTGCTGCTTATGAAGAG-----GCTCTGAGAGCTTCAGTGGCTCC 1199
DB 347 GluCysGlyPheLeuLeuProMetValArgIleLeuAlaValHisThrAlaThrProSer 366
```

```
QY 1200 AACAACTCCCTTGCCCTC-----CATGTGATACAAAGATCTCGACATC 1244
DB 367 AsnArgAspAlaAlaLeuAlaSerThrAlaGlyHisPheHisValArg-----HisIle 384
QY 1245 CCGCTCAGATTTTGCACCTTGTGGCTTAATGAATCTTCATGACACACCTCCCTGACT 1304
DB 385 ProProGlnTyLeuHisIleLeuValAlaAsnGluTyPheHisAspLysHisSerLeuThr 404
QY 1305 GAAATCCGAGACAGTCTTGTGACTTGTGGAGATGTGTTCTTTGTGTGTGTGTGTGTGTGT 1364
DB 405 GluIleAspSerSerLeuLeuAspLeuLeuGlyAspValPhePheValAlaProAlaLeu 424
QY 1365 ATCAAGCTCGATATACAGAGATGCTGTGACCTGTCTACTTCTTAATGATTGCGGAC 1424
DB 425 IleThrAlaArgTyHisAlaArgAspAlaGlyAlaProValTyPheTyGluPheArgHis 444
QY 1425 CCGCTCAGTGTCTTGAAGACAGAGCCGCTTTTGTCAAGCCGACACGCTGATGAA 1484
DB 445 ArgProGlnCysPheGluAspThrLysProAlaPheValLysAlaAspHisAlaAspGlu 464
QY 1485 GTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1544
DB 465 ValArgPheValAlaPheGlyGlyAlaPheLeuLysGlyAspIleValMetPheGluGlyAla 484
QY 1545 ACGAGAGAGAGAAAGTTACTAGCCCGAAGATGAATGAATCTGAGCTTGTGCTGCA 1604
DB 485 ThrGluGluGluLysLeuLeuSerArgLysMetCysTyTrpAlaThrPheAlaArg 504
QY 1605 ACCGGAATCTTAATGGGAAAGACACTGTCTGTGGCCACTTAATACTGATGAGAG 1664
DB 505 ThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpAlaTyRsnLeuThrGluGln 524
QY 1665 TACTTCCAGCTGGAATGAACATGAGCTTGGACAGAGACTCAAGACCGCGGCTGAT 1724
DB 525 TyLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluProAlaGArgAsp 544
QY 1725 TTTTGACACGACCAATCCCC 1745
DB 545 ValTrpValThrGlyTyPro 551

RESULT 9
ADRI9663
ID ADRI9663 standard; protein; 618 AA.
XX
AC ADRI9663;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human drug metabolizing enzyme (DME)-2 protein sequence.
XX
KW drug metabolizing enzyme; DME; cytostatic; immunosuppressive;
KW anti-inflammatory; endocrine; ophthalmological; gastrointestinal;
KW hepatocellular carcinoma; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; endocrine disorder; eye disorder;
KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;
KW human.
XX
OS Homo sapiens.
XX
PN WO200226988-A2.
XX
PD 04-APR-2002.
XX
PF 28-SEP-2001; 2001WO-US030662.
XX
PR 29-SEP-2000; 2000US-0236947P.
PR 06-OCT-2000; 2000US-0238864P.
PR 20-OCT-2000; 2000US-0242323P.
PR 09-NOV-2000; 2000US-0247581P.
PR 16-NOV-2000; 2000US-0249519P.
PR 22-NOV-2000; 2000US-0252834P.
PR 30-NOV-2000; 2000US-0250567P.
XX
```


(INCY-) INCYTE GENOMICS INC.

Azinzai Y, Baughin MR, Borowsky ML, Ding L, Duggan BM; Elliott VS, Gandhi AR, Griffin JA, Hafalia AJA, Ison CH, Khan FA; Lal P, Lee EA, Lu DM, Nguyen DB, Arvizu C, Polichy JL, Rankumar J; Ring HZ, Sanjanwala MS, Tang YT, Tribouley CM, Narinder WK; Walsh KT, Warren BA, Xu Y, Yang J, Yao MG, Yue H; WPI: 2002-362498/39.
N-PsDB; ADR19681.

Human drug metabolizing enzymes, useful in the diagnosis and treatment of disorders associated with aberrant (DME) activity, e.g., cancer and autoimmune disorders.

Claim 1; SEQ ID NO 2; 142pp; English.

This invention relates to novel drug metabolizing enzymes (DME) and the nucleotide sequences which encode them. The invention may be useful for the development of compounds with a cytostatic, immunosuppressive, anti-inflammatory, endocrine, ophthalmological, gastrointestinal or hepatotropic activity acting as an agonist or antagonist of drug metabolizing enzyme activity. The invention may be used in the diagnosis and treatment of disorders associated with decreased or increased expression or activity of drug metabolizing enzymes. Such disorders include cancer, cell proliferative disorders, autoimmune/inflammatory, endocrine, eye, gastrointestinal (including liver disorders) and metabolic disorders. The present sequence is that of a human drug metabolizing enzyme (DME) of the invention. Note: This sequence did not form part of the printed specification but was obtained in electronic format from EPO.

Sequence 618 AA;

Alignment Scores:

Pred. No.:	7.03e-243	Length:	618
Score:	2588.50	Matches:	491
Percent Similarity:	94.62%	Conservative:	1
Best Local Similarity:	94.42%	Mismatches:	5
Query Match:	66.85%	Indels:	23
DB:	5	Gaps:	1

US-10-023-515-1 (1-2158) x ADR19663 (1-618)

186 GGGCCCTTGCCTGAAGGCGCACAGAGAACAACAGAGCTGGATGATTGAGGCAAGCA 245
92 GlyProSerAlaGluGlyProGlnIshGaenThrArgLeuGlyTrpPheGlnIshYrGln 111
246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTTTCCTGGAGTCCCCTTGCTGCT 305
112 ValThrValLeuGlySerProValPProValAsnValPheLeuGlyValPProPheAlaIla 131
306 CCCCCGCTGGGATCCTCGGATTTAGAACCCGCCACTGCAATCGCCCTGGATACTTG 355
132 ProPProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeu 151
366 CGAGAGCCAGCCTCCACCACTTAATTTGTGCGCTCAGAACTCAGAGTGCGCTTAGAT 425
152 ArgGlnIshThrSerTyrrProAsnLeuCysLeuGlnIshSerGluTrpLeuDeuLeuAsp 171
426 CAACATGCTCAGAGGTGATTAACCCGAATTCGAGGTGCAGAGAAGCTGCTTACTTG 485
172 GlnIshMetLeuIshValIshTyrrProLysPheGlyValSerGlnAspCysLeuTyrrLeu 191
486 AACATCTAAGCCGCTGCCCAACGCCCATACAGGCTCCAAGTCCCCGCTTGTGTGTTTC 545
192 AsnIleTyrrAlaPProAlaIshAlaAspPrngIshSerTyrrLeuProValLeuValTrpPhe 211
546 CCAGAGAGGCGCTTCAAGACTGGCTCAGAGCTCATCTTGGATGGGTCGCCGCGGCTGCC 605
212 ProGlyGlyAlaPheLeuYshTrpGlySerAlaSerIlePheAspGlySerAlaLeuAlaIla 231
606 TATGAGAGAGCTGCTGTGTGTGTCGTCAGATACCGGCTAGGAATATTGTGTTCTTACC 665

Db	232	TyrgluarValLeuValValValValGlntrygLeuGlyIlePheGlyPhePheThr	251
Qy	666	ACATGGAGTCAAGCATGCTCCGGAGAACTGGGCTTCAAGGACCAAGTGGCTGTCTCC	725
Db	252	ThrTrpArgGlnIleValIleProGlyAsnTrpIlePheGlyValArgGlnValAlaIleuSer	271
Qy	726	TGGGTCCAGAAAGACATGAGTCTTCCGTGGGAACTCCAGCTGTGGACATCTTGGC	785
Db	272	TrpValGlnValAsnIleGluPhePheGlyGlyIleAspProSerSerValThrIlePheGly	291
Qy	786	GAGTCCGGGGAGGCACTAAGTGTCTTAAGTCTTAACTGTGCTCCCATGGCCAAAGGCTTA	845
Db	292	GlnSerAlaGlyAlaIleSerValSerSerIleuIleuSerProMetAlaGlyGlyLeu	311
Qy	846	TTCCCAAAAGCCATCATGAGAGTGGGGGTGGCCATCATCCCTTACCTGGAGCCCATGAT	905
Db	312	PheIleIleValAlaIleMetGluSerGlyValAlaIleIleProGlyIleGlnIleAlaIleAsp	331
Qy	906	TATGAAAGAGTGAAGACCTGCAAGGTGGTGGCAACATTTCTGTGGTAAACATGGCTCAAC	965
Db	332	TyrglnIleuSerGluAspLeuGlnIleValValAlaIleAspCysGlyValAsnAlaIleSerAsp	351
Qy	966	TCTGAGGCGCTCGTAGGTGGCTTGAAGCAAAACCTCCAAAGAGTGTGTAACCTCAAC	1025
Db	352	SerGlnAlaIleuLeuIleArgCysLeuAlaGlnIleAspSerGlyValIleuIleThrIleuSer	371
Qy	1026	CAGAAACCAAAAGCTCTTTCATCTCGAGTGGTGTATAGTGTCTTCTTCCATAGACCTCTTA	1085
Db	372	GlnIleIleIleuSerPheThrArgValValAlaAspGlyAlaPheIleProAsnGluProIleu	391
Qy	1086	GATCATTTGTCTCAAGAAAGCATTTAAAGCAATTCCTTCCATCATCGAGTCAATTAACAC	1145
Db	392	AspIleuIleuSerGlnIleValIlePheGlyAlaIleProSerIleIleGlyValAlaAsnIleAsp	411
Qy	1146	GAGTGTGGCTTCCGTGCTGCTTATGAAGAGGCTCTGAGTCTCAAGTGGTCCAAACAG	1205
Db	412	GlnCysGlyPheIleuIleuProMet-----	419
Qy	1206	TCCCTGGCCCTCAATGATGACAAACATCTGGACATCCGGCTCAAGTATTTGACATTT	1265
Db	420	-----HisIleProProGlnIleuIleu	428
Qy	1266	GTGGTAAATGAAATCTTCCATGACAAAGACCTCCCTGAATGAAATCCGAGACAGTCTTCTG	1325
Db	429	ValAlaAsnGlnIleuTrpPheIleAspIleuIleSerIleuIleThrGlnIleAspSerIleuIleu	448
Qy	1326	GACTTGGCTTGAAGATGTGTTCTTTGTGGTCCCTGCACTGATCAACAGTCCATATCAACA	1385
Db	449	AspIleuIleuGlyAspValPhePheValValProAlaIleuIleThrAlaIleGlyPheIleArg	468
Qy	1386	GATGTGGTGGACCGTGTACTTCTTAAGAGTTTGGGACCGGCTCAAGTGTTTGAAGAC	1445
Db	469	AspAlaGlyAlaProValIleValIleMetPheGlnIleAlaThrGlnGlnGlnIleuIleu	488
Qy	1446	ACGAGGCGGCTTTTGTCAAAAGCGGACCACTGATGAAGTCCGCTTGTGTTCGGTGT	1505
Db	489	ThrIleuProAlaPheValIleValIleAspIleValIleAspGlnValIleArgPheValPheGlyGly	508
Qy	1506	GCTTCTCTGAAGGGGACATTTATGTTTGAAGAGGCAACGAGAGAGAGAAAGTTACTG	1565
Db	509	AlaPheIleuIleuGlyAspIleValIleMetPheGlnIleAlaThrGlnGlnGlnIleuIleu	528
Qy	1566	AGCGGAAAGATGAAGAAATACTGGGGTCACTTGTCTGGAACCGGGGAATCTTAAGGGGAAC	1625
Db	529	SerArgIleuMetIleuIleuTrpAlaIlePheIleAlaGlnIleGlyAsnProAsnGlyAsn	548
Qy	1626	GACCTGTCTCTGTGGGCACTTATATCTGACTGAGACAGTACTCTCAAGCTGGACTTGAAC	1685
Db	549	AspIleuSerIleuTrpProAlaIleValIleuIleThrGlnGlnIleuIleuAspIleuIleu	568
Qy	1686	ATGAGCTTCGACAGAGACTAAAGAACCGCGGGTGATTTTGGACCAAGCAATGCC	1745

Db 569 MetSerLeuGlyGlnArgLeuGlyGlnProArgArgAspValTrrpValThrGlyTyrPro 588
RESULT 10
ADB64065
ID ADB64065 standard; protein; 469 AA.
XX
AC ADB64065;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human protein encoded by clone BRAMH20021910.
XX
KM Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KM cell regeneration; membrane protein; signal transduction-related protein;
KM transcription-related protein; osteoporosis; neurological disease;
KM cancer; tumour.
XX
OS Homo sapiens.
XX
PN EP1308459-A2.
XX
PD 07-MAY-2003.
XX
PF 28-MAR-2002; 2002EP-00007401.
XX
PR 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Iosogi T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
DR MPI; 2003-450961/43.
DR N-PSDB; ADB62095.
XX
PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
PS Claim 1; Page; 222pp; English.
XX
CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel
CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner, and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a protein of the encoded protein. Note: Some of the
CC specification data for this patent is not represented in the printed
CC European Patent Office.
XX
SQ Sequence 469 AA;
XX
Alignment Scores: 8.14e-230 Length: 469
Pred. No.:

Score: 2454.00 Matches: 467
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.57% Mismatches: 0
Query Match: 63.38% Indels: 0
DB: 7 Gaps: 0
US-10-023-515-1 (1-2158) x ADB64065 (1-469)
QY 432 ATGCTCAAGTGCATTATCCGAAATTCGAGAGTCGAGAAAGCTGCTTACTGAAATC 491
Db 1 MetLeuValValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeuAsnIle 20
QY 492 TTTCGGCTGCTCCACGCGGATACAGGCTTCAGGCTCCCGCTTGTGTGTTCCAGGA 551
Db 21 TyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrrpPheProGly 40
QY 552 GGTCGCTTCAAGACATGCTGCTGCTCCATCTTTATATGGTCCGCTGCTGCTTATGAG 611
Db 41 GlyAlaPheLeuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIleTyrGlu 60
QY 612 GACGTGCTGTGTGTGTGTCAGTACCGGCTAGAAATATTGGATTCTTCACCACTGG 671
Db 61 AspValLeuValValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrTrrp 80
QY 672 GATCAGCATGCTCCGGGGAATCGGCTTCAGGACCAAGGCTGCTGTCTGCTGCTG 721
Db 81 AspGlnHisAlaProGlyAsnTrrpAlaPheLysAspGlnValAlaIleAspTrrpVal 100
QY 732 CAGAGAACATGAGTCTTCGATGCGGAGCCGCTGTCGACATCTTGGCGAGTCC 791
Db 101 GlnLysAsnIleGluPhePheGlyLysAspProSerSerValThrIlePheGlyGluSer 120
QY 792 GCGGAGCCATTAAGTCTTCTTACTGTTTAACTGTTCTTCCATGCGCAAGGCTTATTCAC 851
Db 121 AlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPheHis 140
QY 852 AAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACCTGAGGCCATGATTATGAG 911
Db 141 LysAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyrGlu 160
QY 912 AAGATGAGGACCTGACAGTGTGTGACATTTCTGTGTAAACATGCTCAAGCTGTGAG 971
Db 161 LysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSerGlu 180
QY 972 GCCCTGTGAGGCTGAGGAGCAAAACCTCCAGAGACTGTGACCTCAGCCAGAA 1031
Db 181 AlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrLeuSerGlnLys 200
QY 1032 ACAAGCTTTCACCTGAGTGTGATGCTGCTTCTTCTTATGAGCTCTAGATCTA 1091
Db 201 ThrLysSerPheThrArgValValAlaAspGlyAlaPhePheProAsnGluProLeuAspLeu 220
QY 1092 TTTCCTCAAAAGCATTTTAAAGCAATTCCTTCATCATGAGTCAATTAACAGAGTGT 1151
Db 221 LeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnSerHisGluCys 240
QY 1152 GGCTTCTGCTGCTATGAAAGGAGGCTGAGTCCATGAGGCTCCAAAGTCCCTT 1211
Db 241 GlyPheLeuLeuProMetLysGlnAlaProGlnIleLeuSerGlySerAsnLysSerLeu 260
QY 1212 GCCCTCATCTATACAAACATCTGACATCCGCTCAGTATTTGACCTTGTGCT 1271
Db 261 AlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisLeuValAla 280
QY 1272 AATGAATCTTCAGAGCAAGCATCCCTGACTGAATTCGAGACAGTCTTGTGACTTG 1331
Db 281 AsnGlnTyrPheHisAspLysHisSerLeuThrGlnIleArgAspSerLeuLeuAspLeu 300
QY 1332 CTTGAGATGTGTCTTTGTGTGCTCCCTGACTGATACAGCTCGCATATACAGAGATGCT 1391
Db 301 LeuGlyAspValAlaPhePheValValProAlaLeuIleThrAlaArgTyrHisAlaTyrAla 320
QY 1392 GGTGACCTGTCTACTTCTATGAGATTTTGGCACCGGCTCAGTCTTGAAGACAGAG 1451

DB 321 GYAlaProValrYrPheYrGluPheAArgHLeaRgProGlnCySerPheGluSerThrLys 340
QY 1452 CCGGCTTTTGTGAAGCCCAACCAAGCTGATGAAGTCCCTTTGTCTTCCGTGATGCTTC 1511
DB 341 ProAlaPheValLysAlaSerPheSAlaAspGluValArgPheValPheGlyGlyAlaPhe 360
QY 1512 CTGAAGGGGAGCATTTGTTATGTTGAAAGGACCAAGGAGAGGAAAGTTACTGAGCCGG 1571
DB 361 LeuLySGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuSerArg 380
QY 1572 AAGATGATGAATACTAGGAGCTACCTTTGCTGAAACGGGAAATCTTAATGGAACGACTG 1631
DB 381 LysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProGlnGlyAsnAspLeu 400
QY 1632 TCTCTGTGCGCCGCTTATATCTGACTGAGCAGTACTCTGAGCTGAGCTTGAACATGAGC 1691
DB 401 SerLeuTrpProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMetSer 420
QY 1692 CTCGAGACAGACTCAAGAACCGGGGGGATTTTGGACCAAGCAATCCCGCTGATC 1751
DB 421 LeuGlyGlnArgLeuLysGluProArgValGluPheTrpThrSerThrIleProLeuIle 440
QY 1752 CTGTCTGCTCCGACATGCTCCAGTCTCTTCTTCTTAACTTCTCTCTCTCTC 1811
DB 441 LeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeuLeu 460
QY 1812 CAGCCTTTCTTTTCTTTTGTGCTCT 1838
DB 461 GlnProPhePhePheCysAlaPro 469
RESULT 11
ADP50145
ID ADP50145 standard; protein; 542 AA.
XX
AC ADP50145;
XX
DT 12-FEB-2004 (first entry)
XX
DE Cat cauxin protein SEQ ID NO:2.
XX
KW cat; cauxin; cat kidney disease marker; kidney disease.
XX
OS Felis catus.
XX
PN JP2003250575-A.
XX
PD 09-SEP-2003.
XX
PF 04-MAR-2002; 2002JP-00057908.
XX
PR 04-MAR-2002; 2002JP-00057908.
XX
PA (TOHO-) TOHOKU TECHNORACH KK.
XX
DR MPI: 2004-002277/01.
XX
N-PSDB; ADP50144.
XX
PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
XX and for diagnosing cat kidney disease.
XX
PS Claim 3; SEQ ID NO 2; 33pp; Japanese.

CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.
XX
SQ Sequence 542 AA;
Alignment Scores:
Pred. No.: 5,58e-184 Length: 542
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 51.27% Indels: 2
DB: 8 Gaps: 2
US-10-023-515-1 (1-2158) x ADP50145 (1-542)
QY 186 GGGCCCTTGTGTAAGGGCCAGAGGAACACAGGCTGGATTCAGGGCAAGCA 245
DB 22 GlyProAlaAlaAspAlaProValArgSerThrArgLeuGlyTrpValArgGlyLysGln 41
QY 246 GTCACTGTGCTGGAGAGCCCTGTGCTGAAAGTGTTCCTGGAAGTCCCTTGTGCT 305
DB 42 ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyIleProTyrAlaAla 61
QY 306 CCCCCGCTGGGATCCCTGGATTTACGAACCCGACAGCTGTACCGCCCTGGATTAATTG 365
DB 62 ProProLeuGlyLProLeuArgPheLysGlnProLysProAlaLeuProLysAspPhe 81
QY 366 CGAGAGGACCTCTACCCCTAATTGTGCTCCAGAGCTCAGAGAGGCTGGCTTAGAT 425
DB 82 ArgAsnAlaThrSerTyrProLysLeuCyPheGlnAspLeuGlnTrpLeuValSerTyr 101
QY 426 CAACAATGCTCAAGGTGATTACCCGAATTCGAGAGTGTGAGAGACTGCTTACCTG 485
DB 102 GlnHisValLeuLysValAlaGlyTrpProLysLeuGlnAlaSerGluAspCysLeuTyrLeu 121
QY 486 AACATCTATGCGCCCTGCCCAACCCGATACAGGCTCCAGAGTCCCGCTTGTGCTGCTC 545
DB 122 AsnIleTyrAlaProAlaHisAlaAspAsnGlySerAsnLeuProValMetValTrpPhe 141
QY 546 CCAGAGAGGCTTCAAGCTGCTCAGCTCCATCTTGAAGGGTCCGCTGCTGCTGC 605
DB 142 ProGlyGlyAlaPheLysMetCysSerAlaSerPheAspGlySerAlaLeuAlaAla 161
QY 606 TATGAGAGAGTGTGCTGCTGCTCCAGTACCGGCTTGAAGAAATTTGTTCTTCCACC 665
DB 162 TyrGluAspValLeuIleValIleThrThrGlnTyrArgLeuGlyIlePheGlyPhePheAsp 181
QY 666 ACATGGGATCAGCATGCTCTCCGGGAACTGGGCTTCAAGAGCACAGGTGGCTCTGTCC 725
DB 182 ThrGlyAspGluHisAlaArgIlyAsnTrpAlaLeuLeuAspGlnValAlaIleLeuThr 201
QY 726 TGGGTCCAGAAAGAAATGATGATTTCTTGGTGGGAAACCCAGTCTGTGACATCTTGGC 785
DB 202 TrpValArgAspAsnIleGluPhePheGlyGlyAspProArgSerValThrIlePheGly 221
QY 786 GAGTCCGGGGAGGACATTAAGTGTTCATGCTTATAGTGTCTCCCATGGCCAAAGCTTA 845
DB 222 GluSerAlaGlyAlaIleSerValSerLeuIleLeuSerProIleAlaAsnGlyLeu 241
QY 846 TTCACAAAGCCATCATGAGAGAGTGGGGTGGCCATCATTCCTTACCTG--GAGGCCAT 902
DB 242 PheHisLysAlaIleMetGluSerGlyValAlaIleLeuProLeuLeuMetArgProPro 261
QY 903 GATTATGGAAGAGAGAGAGCTGAGAGTGTGTGGAACATTTCTGTGTTAAACAATGCTCA 962
DB 262 GlyAspGluArgLysLysAspLeuGlnValLeuAlaArgIleCysGlyCysHisAlaSer 281
QY 963 GACTGTGAGGCTCTGATGAGTGTGCTGAGGACAAACCTCCAAAGAGCTGTGACCTTC 1022


```

Oy 726 TGGGTCAGAGAACATGAGTTCTTCGGTGGGAGCCAGCTCTGTGACCATTTGGC 785
Db 202 TTPVAlaIgaApheNIIleGIuPheheGlyIaPheProAgsSerValThrlIleheGly 221
Oy 786 GAGTCGGGGAGCCATAAGTGTTCAGTCTTAATAGTCTCCAGTGGCCAAAGCTTA 845
Db 222 GluSerAlaIgaIaIleSerValSerSerIleuIleuSerProIAlaIaIaIleu 241
Oy 846 TTCCCAAAAGCCATATGAGAGTGGGGTGGCCATCATCCCTTACTCTG---GAGGCCAT 902
Db 242 PheHsIyValaIleMeGluSerIyValaIaIleuProIleuMeuAArgProPro 261
Oy 903 GATTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
Db 262 GlyApheIuIaIyIaIleuIaIleuIaIleuIaIleuIaIleuIaIleuIaIleu 281
Oy 963 GACTCTGAGGCGCTGCTGAGAGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1022
Db 282 ApsSerAlaIaIleuIeGluIeGluIeGluIeGluIeGluIeGluIeGluIeGluIe 301
Oy 1023 AGCCAGAAACAAAGTCTTCTCAGTGGTGTATGTGCTTCTTCTTCTTATAGCTT 1082
Db 302 SerIySlyLeuThrPheSerIleProValIleApsPhePheProApsIuPro 321
Oy 1083 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCATTCCTTCATCATGAGTGGATTAAC 1142
Db 322 ValAlaIeIuThrIuIaIleuIaIleuIaIleuIaIleuIaIleuIaIleuIaIleu 341
Oy 1143 CACGAGTGTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1202
Db 342 HisGIuIyValaIlePheIeIu---SerThrGIuIePheSerGIuIleIeGluIyGIeSerHs 360
Oy 1203 AAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1262
Db 361 ArgSerIeIuIaIleuIyIeIuValHsThrPheIeIuIaIleProThGIuIyIeHs 380
Oy 1263 CTTGGTGGTATGATATCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1322
Db 381 LeuValAlaApsHsIyIyPheIyIaIleuIaIleuIaIleuIaIleuIaIleuIaIleu 400
Oy 1323 CTGAGACTGTGCTGAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1382
Db 401 LeuApsIeIuIeGluIaIleuIaIleuIaIleuIaIleuIaIleuIaIleuIaIleu 420
Oy 1383 AGAGATGTGTGTGCACTGTCTTCTTATGAGTTGGAGAGAGAGAGAGAGAGAGAG 1442
Db 421 ArgApsAlaIgaIaIleProValIyIyPheIyIaIleuIaIleuIaIleuIaIleu 440
Oy 1443 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1502
Db 441 ApsThrApsProAlaPheValIaIleuIaIleuIaIleuIaIleuIaIleuIaIleu 460
Oy 1503 GGTGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1562
Db 461 GIyIaIlePheIeIuIyGIyApsIleValIeIuIePheGIuIyIaIleThGIuIyIeIu 480
Oy 1563 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1622
Db 481 LeuSerAArgIyIaIleuIaIleuIaIleuIaIleuIaIleuIaIleuIaIleu 500
Oy 1623 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1682
Db 501 GIuIyIaIlePheIeIuIyGIyApsIleValIeIuIePheGIuIyIaIleThGIuIyIeIu 520
Oy 1683 AAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1742
Db 521 SerValSerValGIyGIuIyIeIuIyIeIuIyIeIuIyIeIuIyIeIuIyIeIuIyIeIu 540

```

```

AC ABP61006;
XX
XX 10-SEP-2002 (first entry)
DT
XX
XX Novel human protein. SEQ ID 93.
DE
XX
XX Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian;
XX neurotropic; neuroprotective; immunosuppressive; haemostatic;
XX antiinflammatory; cardiant; antitumor; virucide; antihypoid;
XX wound healing disorders; atherosclerosis; Parkinson's disease;
XX Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
XX inflammation; neoplastic disease; nervous system disorder;
XX cardiovascular disorders; pancreatitis; respiratory disorder;
XX hyperproliferation; systemic autoimmune disease; hyper-immunity;
XX developmental abnormality; gastrointestinal ulceration; neuropathy;
XX haematological disease; metabolic disease; sperm dysfunction;
XX thyroid disorder; hypothyroidism; brain damage; colitis;
XX cone photo- transduction deficiency; neurological disease; stroke;
XX angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
XX trachea; thymus; lymph node; muscular system; obesity; anorexia;
XX growth abnormality; precocious puberty.
XX
XX Homo sapiens.
OS
XX
XX WO0205105-A1.
XX
XX 27-JUN-2002.
XX
XX 17-DEC-2001; 2001WO-US049232.
XX
XX 19-DEC-2000; 2000US-0256710P.
XX 20-DEC-2000; 2000US-0257048P.
XX 09-JAN-2001; 2001US-0260482P.
XX 30-JAN-2001; 2001US-0264922P.
XX 06-FEB-2001; 2001US-0266797P.
XX 19-MAR-2001; 2001US-0276988P.
XX 04-APR-2001; 2001US-0281535P.
XX 08-MAY-2001; 2001US-0289622P.
XX
XX (SMK ) SMITHKLINE BEECHAM CORP.
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KP, Lai Y;
XX Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX WPI; 2002-508784/54.
XX N-PSDB; ABQ86171.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
XX or treating various diseases e.g. cancer, wounds, atherosclerosis,
XX Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX Claim 1(a); Page 315; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
XX which allow it to be secreted extracellularly or membrane associated. The
XX activity of polypeptides of the invention may be described as,
XX cytostatic, vulnary, antiarteriosclerotic, antiparkinsonian, neurotropic,
XX neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
XX cardiant, antitumor, virucide, antihypoid, cerebroprotective, anorectic,
XX and metabolic. Polypeptides and polynucleotides of the invention are
XX useful in the treatment, or as a vaccine in the prevention of, cancer,
XX wound healing disorders, infection, atherosclerosis, Parkinson's disease
XX and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
XX inflammation, neoplastic diseases, nervous system related disorders and
XX cardiovascular disorders, pancreatitis, respiratory disorder,
XX hyperproliferation, systemic autoimmune disease, hyper-immunity,
XX developmental abnormality, gastrointestinal ulceration, neuropathy,
XX haematological diseases, metabolic diseases, sperm dysfunction, thyroid
XX disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
XX transduction deficiency, neurological diseases, stroke, angiogenesis,
XX

```

CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC tracheas, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABP60965-ABP61019 represent novel human
 CC proteins of the invention

XX Sequence 356 AA;

Alignment Scores:

Pred. No.:	4,556-164	Length:	356
Score:	1780.00	Matches:	340
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	45.97%	Indels:	0
DB:	5	Gaps:	0

US-10-023-515-1 (1-2158) x ABP61006 (1-356)

QY 819 ATACTGCTCCCATGGCCAAAGGCTTATTCACAAAGCCATCATGGAGTGGGCTGCC 878
 DB 17 ILeuSerProMetAlaLysGlyLeuPheHISlyeAlaIleMetGluSerGlyValAla 36
 QY 879 ATCATCCCTTACCGGAGGCCCATGATTAGAAAGATGAGGACCTGCGAGGTGGTGA 938
 DB 37 ILeileProIlyLeuGluAlaHISAspTYGILySserGluAspLeuGlnValAla 56
 QY 939 CATTTGTGTGATACAAATGCGTCAAGCTGAGGCCCTGCTGAGGTGCGCTGAGGACAA 998
 DB 57 HIsPheCyGlyAsnAsnAlaSerAspSerGluAlaLeuLeuHISGlySerLeuArgHIS 76
 QY 999 CCTTCAAGGAGGCTGACCCCTGACCCGAGGAAACAAAGCTTTTCACTGAGGTGGTAT 1058
 DB 77 ProSerIyGlyLeuLeuThrLeuSerGlnIySethIySserPheThrArgValAlaAsp 96
 QY 1059 GGTGCTTTCTTCTATGAGGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAAT 1118
 DB 97 GILyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnIyAlaPheIyAlaIle 116
 QY 1119 CCTTCATATGAGGATGCAATTAACAACAAGGTGGCTTCTGCTGCTTATGAGAGGCT 1178
 DB 117 ProSerIleIleGlyValAlaAsnAsnISGILySgIyPheLeuLeuProMetIyGlnAla 136
 QY 1179 CCTGAGATCTCTGAGTGGCTCAACAAGTCCCTTGGCCCTCATGATCAAAACATTCCTG 1238
 DB 137 ProGluIleLeuSerGlySerAsnIySserIyAlaLeuHISleuIleGlnAsnIleLeu 156
 QY 1239 CACATCCCGGCTCAGATTTTGCACCTTGTGCTAATGAAATCTTCATGACAGCACTCC 1298
 DB 157 HIsIleProProGlnIyIyLeuHISleuValAlaAsnGILyTYrPheHISAspIyHISer 176
 QY 1299 CTGACTGAATCCGAGACAGTCTTGTGACTTGGTGGAGATGTGTTCTTTGTGGTCCCT 1358
 DB 177 LeuThrGluIleHISAspSerIyLeuAspLeuLeuGILyAspValAlaPhePheValAlaPro 196
 QY 1359 GCACTGATCAAGCTCATATCACAGAGATCTGGTGCACCTGTCTACTTATGAGATT 1418
 DB 197 AlaleuIleThrAlaArgTYrHISArgAspAlaGILyAlaProValTYrPheTYGlnPhe 216
 QY 1419 CGGACCGGCGCTCGTGTGTTTGAAGACACGAGCGGCTTTTGTCAAGCCGACCAAGCT 1478
 DB 217 ArgHISAspProGlnCyAspHISAspThrIyAspProAlaPheValIyAlaAspHISAla 236
 QY 1479 GATGAAGTCCGCTTTGTGTTGCTGAGTGCCTTCTGAAAGGGGAGCATTTGTTATGTTGA 1538
 DB 237 AspGILyAlaArgPheValPheGILyGILyAlaPheLeuIySgIyAspIleValISerPheGln 256
 QY 1539 GAGGCCACGAGAGGAGGAAAGTACTGAGCCGGAAGATGATGAATATCTGGGCTACCTTT 1598
 DB 257 GILyAlaThrGluGluGluIyIyLeuLeuSerArgIySethMetCylSgTYrTPAlaThrPhe 276
 QY 1599 GCTTGAACCGGGAATCTTAATGGAGACAGCTGTCTGTGGGCGAGTTTATCTGAGCT 1658
 DB 277 AlaArgThrGILyAsnProAsnGILyAsnAspLeuSerIyProAlaTYrAsnLeuThr 296

QY 1659 GAGAGTACTCTCAGCTGAGTCAATTAAGACCTCGGACAGAGACTCAAGAACCGGG 1718
 DB 297 GILyGlnTYrLeuGlnLeuAspLeuAsnMetSerIyGILyGlnArgLeuIySgILyProArg 316
 QY 1719 GTGAGTTTGTGACCAAGACCAATCCCTGATTCCTGTGCTCGGACATGCTCCACAGT 1778
 DB 317 ValAspPheThrIyPheSerThIleProLeuIleLeuSerAlaSerAspMetLeuHISer 336
 QY 1779 CCTTTTCTTCTTAACTTCTGCTGCTGCTGAGGCTTTCTTTTCTTTGTGGCTCCT 1838
 DB 337 ProLeuSerSerIyLeuThrPheLeuSerIyLeuGlnProPhePhePheCyAlaPro 356
 RESULT 14
 AAE20911
 ID AAE20911 standard; protein; 561 AA.
 AC XX
 AC AAE20911;
 XX XX
 DT 01-JUL-2002 (first entry)
 XX XX
 DE Rat carboxylesterase-like enzyme protein.
 XX XX
 KW Rat; carboxylesterase-like enzyme; organophosphorus intoxication;
 KW osteoporotic; gene therapy; osteoporosis; antisense therapy; cytostatic;
 KW detoxifying agent; Paget's disease; bone implant degradation; cancer;
 KW dental implant; enzyme; gene expression.
 XX XX
 OS Rattus sp.
 XX XX
 PN WO200206454-A2.
 XX XX
 PD 24-JAN-2002.
 XX XX
 PE 10-JUL-2001; 2001WO-EP007919.
 XX XX
 PR 17-JUL-2000; 2000US-0218564P.
 XX XX
 PA (FARB) BAYER AG.
 XX XX
 PI Xiao Y;
 XX XX
 DR WPI; 2002-195808/25.
 XX XX
 PT Novel human carboxylesterase-like enzyme polypeptide, regulators of which
 PT are useful for preventing and treating organophosphorus intoxication,
 XX cancer and osteoporosis.
 PS Disclosure, Fig 8; 92pp; English.
 XX XX
 CC The invention relates to a purified human carboxylesterase-like enzyme
 CC polypeptide. Carboxylesterase-like enzyme and its DNA are useful for
 CC screening for agents which decrease or modulate the activity of
 CC carboxylesterase-like enzyme polypeptide. Carboxylesterase-like enzyme is
 CC useful for treating a carboxylesterase-like enzyme dysfunction related
 CC disease, such as organophosphorus intoxication, cancer and osteoporosis.
 CC Compounds that increase the ability of human carboxylesterase-like enzyme
 CC to bind to organophosphorus compounds are useful as detoxifying agents.
 CC Carboxylesterase-like enzyme agonists and antagonists are useful for
 CC treating osteoporosis. Paget's disease and degradation of bone implants,
 CC particularly dental implants. Carboxylesterase-like enzyme is useful in
 CC diagnostic assays for detecting diseases and abnormalities or
 CC susceptibility to diseases or abnormalities related to the presence of
 CC mutations in the nucleic acid sequences which encode the enzyme. The
 CC coding sequence of carboxylesterase-like enzyme polynucleotide is useful
 CC in gene therapy and for generating antisense oligonucleotides or
 CC ribozymes which specifically bind to mRNA transcribed from
 CC carboxylesterase-like enzyme DNA. These antisense oligonucleotides are
 CC useful for modulating carboxylesterase-like enzyme gene expression. The
 CC present sequence is rat carboxylesterase-like enzyme protein
 XX
 XX Sequence 561 AA;
 SQ

Alignment Scores:

Pred. No.: 2,04e-113 Length: 561
 Score: 1262.50 Matches: 275
 Percent Similarity: 62.41% Conservative: 77
 Best Local Similarity: 48.76% Mismatches: 179
 Query Match: 32.61% Indels: 33
 DB: 5 Gaps: 13

US-10-023-515-1 (1-2158) x AAB20911 (1-561)

```

OY 96 ATGCCAAGGGA---CTTACTTCATCTGCTTCAAGAGCTGCTTTTCTGATTTCCAG 152
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1 MetProAGuAnGlnLeuHisSerTPrLeuAsnAlaValLeuPheGlyLeuLeuLeu--- 19
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 153 CCCCTGTTGGGACAGAGATGGGGAAGAAACCTGGGCTTTCCTGACAGGCCACAGAG 212
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 20 --LeuLeuIleHisValGln--GlyGlnAerSerProGluSer---SerProIleAerG 36
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 213 AACACAGGCTGGAGTGAATTCAGGAGCAAGATCACTGTGCTGGAGACCTGTGCT 272
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 37 ThrThrHisThrGlyGlnValGlnGlyValLeuAerHisValLysAerThrLysAlaGly 56
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 273 GTGAACGTGTCTTCGAGACTCCCTTTGCTGCTCCCGCTGGAGATCCCTGCGATTACG 332
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 57 ValHisThrPheLeuGlyLysProPheAlaLysProValGlyProLeuAerPheAla 76
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 333 AACCCGACGCTGCATCGCCCTGGAGTAACTTGCAGAAAGCCATCTCTACCTTAATTG 392
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 77 ProProGluProProGluProTTrPheSerGlyValArgAerAlaThrSerGlnProAlaMet 96
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 393 TGCCTCCAGAACTCAAGATGGCTG-----CTCTAGATCAACACATGCTCAAG 440
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 97 CysLeuGlnAerHisLeuAerHisLeuAerHisLeuValGlyLeuAerHisMetLysMetLys 115
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 441 GTGCATTACCCGAAATTCGAGAGTGCAGAAAGCTGCTTCACTTGAACATTAAGGCT 500
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 116 -----LeuSerSerIleSerMetSerGlnAerCysLeuLysLeuAerHisValLysAlaPro 133
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 501 GCCCAGCCGAGTACAGAGCTCCCAAGCTCCCGCTTTCGCTGCTCCCAAGAGTGGCTTC 560
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 134 AlaHisValArgGlyGlySerHisLeuProValMetValTrpIleHisGlyGlyAlaLeu 153
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 561 AAGATGCTGACGCTCCATCTTTGATGGGCTCGGCTGGCTGCTGCTTACAGAGAGCTGCTG 620
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 154 ValValGlyMetHisLeuSerMetLysAerHisLeuThrValAerHisLeuVal 173
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 621 GTTGTGCTGCTCAGTACCGGCTAGAGATATTTGCTTCTTCAACCATGAGTACAGAT 680
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 174 ValValThrIleGlnLysLeuGlyValLeuGlyPhePheSerThrGlyAerHis 193
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 661 GCTCCGGGGAATGGGACCTTCAAGACAGAGCTGCTGCTGCTGGGCTCCAGAAAGAC 740
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 194 AlaArgGlyAerTrpGlyLysLeuAerHisValAlaLeuAerHisValAlaGlnGlnAer 213
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 741 ATGAGTCTTCTGAGTGGGAGCCCAAGCTGTGACATCTTTGGGAGTCCGCGGAGAGCC 800
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 214 IleAlaHisPheGlyGlyAerHisProHisAerHisValIlePheGlyGlySerHisAlaGly 233
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 801 ATAGATGTTTCTAGTCTTATACGTCTCCCATGAGCCAAAGGCTTATTCACAAAGCCATC 860
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 234 ThrSerValSerSerHisValIleSerProMetSerGlnGlyLeuPheHisGlyAlaIle 253
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 861 ATGAGAGTGGGCTGGCCATCATCTCTTACCTGAGAGGCCCATGATTAAGAAAGTGAAG 920
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 254 MetGlnSerGlyValAlaLeuLeuProAerHisLeuSer-----GlnHisSerGln 270
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 921 GACCTGACG--GTGGTTGACATTTCTGTGGAACAATGGGACAGACTCTGAGGCGCTG 977
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 271 ThrValSerThrThrValAlaLysLeuSerGlyCysGlyAlaMetSerHisValAlaLeu 290
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 978 CTGAGAGTCTTGAAGCAAAACCTTCCAGAGAGCTGTACCTCAGCCAGAAACCAAG 1037
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 291 ValArgCysLeuArgAlaLysSerGlyAlaGlnIleLeuValIleAerHisValPheLys 310
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

OY 1038 TCTTTCACTCGAGTGGTATGCTGCTTTCTTTCTATATGAGCTTGAATGCTAT 1097
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 311 MetIleProAlaValValAerHisLeuPheLeuProAerHisLeuProGlyLeuAla 330
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1098 CAGAAAGCATTTAAAGCATTTCTTCCATCATCGAGTCAATGAACAGAGTGGCTTC 1157
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 331 SerGlnAerPheHisLeuProAlaProSerIleIleGlyValAerHisThrAerHisLysCys 350
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1158 CTGCTGCTTATG-----AAGAGGCTCTCGAGATCTTCAGT 1193
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 351 ThrIleProMetValMetGlyThrAlaGlnIleLysGln-----LeuSer 366
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1194 GGCCTCAACAATGCTCCCTGCTCCATCATGATGAACAAACATCGTCAATCCGCTGAG 1253
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 367 ArgGlnAerHisLeuGlnAlaValLeuAerHisThrAlaAlaGlnMetLeuProProGln 386
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1254 TATTTGACCTTGTGGCTAATGAATCTTCCATGACAAAGCACTCCCTGACTGAATCCGA 1313
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 387 CysGlyAerHisLeuMetGlnGlyLysMetGlyAerHisThrAerAerSerGlnHisLeuGln 406
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1314 GACAGCTTTCGAGCTGCTTGGAGATGCTTTCTTTGCTGCTCCCTGACATGATCAGCT 1373
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 407 IleGlnLysThrGlnMetMetGlyAerPheLeuPheValIleProAlaLeuGlnValAla 426
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1374 CGATATACAGAGATGCTGTCACCTGTCTTCTTGAAGTTTGGGACCGGCTCAG 1433
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 427 HisPheGlnArg--SerHisAerProAlaThrPheGlyGlnPheGlnHisValAerProSer 445
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1434 TGCCTTGAAGACAGAAAGCTGCTTTTGTCAAGGCGGACAGCTGATGAAGTCCGCTTT 1493
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 446 TyrPheLysAerHisValArgProHisValLysValAerHisAlaAerHisValAerProPhe 465
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1494 GTTGTGCTGCTGCTCTTCAAGGAGCAATTTGATTTTGAAGAGCCAGAGAGAG 1553
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 466 ValPhe--GlySerPhePheSerGlyMetLysLeuAerPhe-----ThrGlnGln 481
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1554 GAGAGTCTTGAAGCGGAGAGTGAATGAATCTGAGGCTTCACTTTCGAAACCGGAGAT 1613
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 482 GlnAerHisLeuSerHisArgHisMetCysLysTrpAlaAerHisValAerHisGlnGlnAer 501
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1614 CCTAATGGGAGAGCACTGCTGCTGCTGAGCTTATCTGATGATGAGAGTACCTCCAG 1673
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 502 ProAerHisLeuGlnGlyLeuProGlyTrpProAlaLeuAerHisAerHisLeuHisLeuGln 521
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1674 CTGAGACTTGAATGAGCTCGAGACAGATCTGAAGAAACCGGAGTGAATTTTGAAG 1733
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 522 LeuAerThrHisLeuProAlaValAerHisValAerHisValAerHisLeuHisLeuHis 541
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1734 AGCACCATCCCC 1745
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 542 LysThrLeuPro 545
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
AAB58981
ID AAB58981 standard, protein, 549 AA.
XX
AC AAB58981;
XX
DT 27-MAR-2001 (first entry)
XX
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 689.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neoplastic; neuroprotective; antiviral; anti-allergic; hepatotropic;
KW antidiabetic; anti-inflammatory; antitumor; antitumor; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune hemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.
XX
OS Homo sapiens.
  
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen, Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: June 15, 2005, 06:53:49 ; Search time 38.97 Seconds
(without alignments)
10656.177 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 3872
Sequence: 1 ccacgcgtccgaacacagc.....aaaaaaaaaaaaaaaaa 2158

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2.1/USPTO_epool/6664091/runat_14062005_133459_9346/app_query.fasta_1.4238
-DB=FIR 79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091 @CGN 1 1 101 @runat_14062005_133459_9346 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 79: +
1: pir1: +
2: pir2: +
3: pir3: +
4: pir4: +

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1298.5	33.5	557	2	A47162
2	1231.5	31.8	532	2	A34329
3	1230	31.8	559	1	JC5408
4	1199	31.0	561	2	S47855
5	1186.5	30.6	554	1	S34607
6	1163	30.0	554	2	A39060
7	1162.5	30.0	566	2	S19307
8	1144	29.5	549	2	JX0054
9	1140	29.4	567	1	A41010
10	1139	29.4	561	2	S62788
11	1138.5	29.4	562	2	A55281
12	1136	29.3	540	2	A31584
13	1129.5	29.2	561	2	S71597
14	1123	29.0	561	2	JC2447

15	1121.5	29.0	565	2	S10367	carboxylesterase (
16	1034.5	26.7	539	2	A29923	carboxylesterase (
17	777	20.1	596	1	ACR9E	acetylcholinestera
18	764.5	19.7	614	2	A39256	acetylcholinestera
19	763	19.7	614	2	JH0314	acetylcholinestera
20	756	19.5	599	1	A38868	acetylcholinestera
21	754	19.5	614	2	JH0811	acetylcholinestera
22	752	19.4	602	1	ACHU	cholinesterase (EC
23	740	19.1	603	2	S70849	cholinesterase (EC
24	732	18.9	581	2	C39768	cholinesterase (EC
25	726.5	18.8	584	2	S48724	acetylcholinestera
26	710.5	18.3	583	2	S10712	acetylcholinestera
27	692.5	17.9	620	2	A54413	acetylcholinestera
28	669.5	17.3	745	2	S13586	triacylglycerol 11
29	658.5	17.0	489	2	B65680	para-nitrobenzyl e
30	643.5	16.6	597	2	A33668	sterol esterase (E
31	633.5	16.4	599	2	A57701	sterol esterase (E
32	627.5	16.2	664	2	UC7990	acetylcholinestera
33	618	16.0	612	2	A34967	sterol esterase (E
34	594.5	15.4	691	2	JE0150	acetylcholinestera
35	593	15.3	550	1	A34576	crystal protein pr
36	588	15.2	767	2	S47639	acetylcholinestera
37	581.5	15.0	746	2	A25363	acetylcholinestera
38	575	14.9	640	2	S51043	carboxylesterase (
39	572.5	14.8	502	2	T37254	acetylcholinestera
40	570.5	14.7	602	2	S66236	acetylcholinestera
41	569.5	14.7	534	2	T31783	hypothetical prote
42	569.5	14.7	557	2	A56690	esterase (Caenorh
43	562.5	14.5	545	2	A69046	protein B0238.1 [1
44	557.5	14.4	593	1	S25062	triacylglycerol 11
45	553	14.3	956	2	A56920	gliotactin precurs

ALIGNMENTS

RESULT 1

A47162

Cholinesterase B (BC 3. - - -) precursor - mallard

C/Species: Anas platyrhynchos (mallard)

C/Date: 05-May-1995 #sequence #revision 05-May-1995 #text_change 09-Jul-2004

C/Accession: A47162

R./Hwang, C.S.; Kolatukudy, P.E.

J. Biol. Chem. 268, 14278-14284, 1993

A./Title: Molecular cloning and sequencing of cholinesterase B cDNA and stimulation of expe

A./Reference number: A47162; MUID:93300823; PMID:8314791

A./Accession: A47162

A./Status: preliminary

A./Molecule type: mRNA

A./Residues: 1-557 <HMA>

A./Cross-references: UNIPROT:Q04791; GB:L05493; NID:G213100; PIDN:AAA49223.1; PID:G213101

C./Superfamily: cholinesterase; cholinesterase homology

C./Keywords: hydrolase

F./56-545/Domain: cholinesterase homology <CHE>

Alignment Scores:

Pred. No.: 7.1e-91 Length: 557
Score: 1298.50 Matches: 265
Percent Similarity: 62.77% Conservative: 79
Best Local Similarity: 48.36% Mismatches: 177
Query Match: 33.54% Indels: 27
DB: 2 Gaps: 8

US-10-023-515-1 (1-2158) x A47162 (1-557)

QY	183	ACTGGGCTTGTGTTGAAGGCGACAGAGAACACACAGCTGGATGATTCAGGCAAG	242
DB	23	ThrglyglnlyslalaglglinProgluValValThrAenTyrglySerValArglylyr	42
QY	243	CAAGTCATGTCGTGGGAAGCCCTGCTGTAACGTTCTCTGGAGTCCCTTGGCT	302
DB	43	GlnVallylValAsnalalaglulargSerValenValPheleuglyleuprotheala	62
QY	303	GCTCCCGCGGTGGATCCCTGGATTTACGAACCGGACGCTGATCCCTGGATTAAC	362

```

Db      63  LysProValIglyProLeuAhrgheserGluProGlnProProGlnProTrrlyGly 82
Qy      363  TTGGAGAAAGCCACCTCTTAACCTTAATTGGCTCCGAACTCAGAGCTGCTCTTA 422
Db      83  ValAgaAraAlaAserTyrProMetCysLeuGlnAhrlys-----ValLeu 99
Qy      423  GATCAACATGCTC-----AAGGTGATTAACCGGAATTCCGAGTGTCA 467
Db      100  GlyGlnTyrLeuSerAraAlaIleThrAsnArglySerGlyValAhrGLeuGlnTrrSer 119
Qy      468  GAAAGCTCTCTTACCTTAACATCTATGCGCTCCGACCGCCATACAGGCTCCAGCTC 527
Db      120  GlnAhrCysLeuTyrLeuAhrAhnValTyrThrProValSerThrGlnGlnGlnGlyLeu 139
Qy      528  CCCGCTTGGTGTGATTCGAGAGAGGCTTCAAGACTGCTCAGAGCTCTCATCTTGTAT 587
Db      140  ProValAhrValTrrPrrIleAhrGlyGlyLeuValSerGlyAlaAhrSerTyrAhr 159
Qy      588  GGGTCCCGCTGCTGCTTATGAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 647
Db      160  GlySerAlaLeuAlaAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 179
Qy      648  AATTTGGTTCTTACACCAATGAGATCAGATGCTCCGGGAACTGGCTTCAAGAC 707
Db      180  IleAhrGlyTyrAhrSerThrGlyAhrPrrAhrAhrAhrAhrAhrAhrAhrAhrAhr 199
Qy      708  CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 767
Db      200  GlnValAlaAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 219
Qy      768  TCTGTACCACTTTGGCGAGCTCCGGGAGCCATAGTGTCTTATGCTTACTGTCT 827
Db      220  SerValThrIleAhrGlyGlnSerAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 239
Qy      828  CCCATGCGCAAGAGCTTATTCAGAAAGCATATGAGAGAGGAGGCTGCGCATC---ATC 884
Db      240  ProLeuAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 259
Qy      885  CCTTACCTGAGAGCCATATTAAGAAAGTGAAGAGCTGAGGCTGCTGCTGCTGCTGCT 944
Db      260  LeuPrrAhrGln-----GlnProGlnGlnGlnAhrAhrAhrAhrAhrAhrAhrAhr 275
Qy      945  TGTGTAAACATGCGTCAAGTCTGAGAGCTGCTGAGTGTCTGAGCAAAACCTTCC 1004
Db      276  AlaGlyCysGlnAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 295
Qy      1005  AAGAGAGCTG-----CTGACCTTCAAGCCAGAAAGAAAGTCTTTCATCTGAGTGTGAT 1058
Db      296  AlaGlnMetGlnGlnIleThrLeuAhrMetProPrrMetPrrIleSerAhrAhrAhr 315
Qy      1059  GGTGCTTCTTCTTCAATGAGCTCTAATCTATTTGCTCAAGAAAGCATTTAAAGCAAT 1118
Db      316  GlyValAhrPrrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 335
Qy      1119  CTTTCCATCATCGAGTCAATTAACAGAGTGTGCTTCCGCTGCTGCTGCTGCTGCTGCT 1175
Db      336  ProTrrIleIleGlyValAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 355
Qy      1176  GCTCTGAGATCTCTGAGTGTCCAAAGATCCCTTGTCCATCTGATCAAAAGATC 1235
Db      356  PrrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 375
Qy      1236  CTGCAACATC-----CCGCTCAAGTATTTGACCTTGTGGCTCAATGATATAC 1280
Db      376  LeuAlaAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 395
Qy      1281  TTCCATACAAAGACCTCCCTGATGAAATCCGAGAGTGTGCTTGTGAGCTTGTGAGAT 1340
Db      396  IleGlyValAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 415
Qy      1341  GTGTTCTTTGTGTGCTCCCTGACATATCAAGCTGTGATATCAAGAGTGTGTGACCT 1400

```

```

Db      416  ProLeuPrrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 435
Qy      1401  GTTACTTCTATGAGTTCGAGCCGGCTCAGTCTTGTGAAGACAGACCGGCTTTT 1460
Db      436  ValTrrPrrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 455
Qy      1461  GTCAAGCCGACCCAGTATGAGTTCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1520
Db      456  ValAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 475
Qy      1521  GACATTTGATTTGGAAGAGCCAGAGAGAGAAATTATGAGCCGAGATGATG 1580
Db      476  Ahr-----AlaThrGlnGlnGlnAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 489
Qy      1581  AATATCGGCTACCTTGTCTGAAACCGGATCTTAATGGGAACGCTGTCTGTGG 1640
Db      490  LysTrrTrrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 509
Qy      1641  CCAAGTTTAAATCTGATCAGACAGTCTTCAAGCTTGAATGAATGAGCTTGGACAG 1700
Db      510  ProGlnTrrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 529
Qy      1701  AGACTCAAGAAACCGGCTGATTTTGAACAGACATCCCTGATCTGTCTGCC 1760
Db      530  LysLeuAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 549
Qy      1761  TCCGACATGCTCCACAGTCTTT 1784
Db      550  ArgAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 557

RESULT 2
A34329
60K esterase (EC 3.1.1.-) isoform 2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 09-Jul-2004
C:Accession: A34329
R:Ozols, J.
J: Biol. Chem. 264, 12533-12545, 1989
A:Title: Isolation, properties, and the complete amino acid sequence of a second form of
A:Reference number: A34329; MUID:89308686; PMID:2745458
A:Accession: A34329
A:Status: preliminary
A:Statue: preliminary
A:Molecule type: protein
A:Residues: 1-532 <OZO>
A:Cross-references: UNIPROT:P14943
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
F:32-517/Domain: cholinesterase homology <CHE>
F:201/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 9,13e-86 Length: 532
Score: 1231.50 Matches: 253
Percent Similarity: 62.19% Conservative: 76
Best Local Similarity: 47.83% Mismatches: 167
Query Match: 31.81% Indels: 33
DB: 2 Gaps: 7

US-10-023-515-1 (1-2158) x A34329 (1-532)
Qy      204  CCAAGAGAAACAAGAGCTGATGATTCAGGCAAGTCACTGTCTGGAGAC 263
Db      6  ProIleAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 25
Qy      264  CCGTGTGCTGAGAGTGTCTGAGAGTCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 323
Db      26  AhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 45
Qy      324  CGATTTAAGAAACCGGACCTGATCGCTGAGATACTGGAGAGACCAAGCTTCAAC 383
Db      46  ArgPrrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhrAhr 65
Qy      384  CCTAATTGTGCTCCAGAACTCAAGAGTGTGCTTGAATCAACATGCTCAAGGTG 443

```



```

Db      66 ProAlaMeCyLeuGlnAen-----LeuAlaIleMeAspGlnAerValIleuLeu 83
Qy      444 CATTAC-----CCGAAATTCGAGTGTGAGAAAGCTGCTTACTTGAACATCTATGG 497
Db      84 HbPheThrProPheSerIleProMetSerGlnAerCyLeuTyLeuAenIleTySer 103
Qy      498 CTTGCGCCAGCGGATACAGGCTCCAAAGCTCCCGCTTGTGTGGTTCACAGAGGTGC 557
Db      104 ProAlaHbAlaArgGlnAerGlnAerLeuProValMetValTrrIleHbGlyGly 123
Qy      558 TTCAAGACTGAGCTCAAGCTTCATTTGATGGCTCCGCTGAGCTTGAAGAGCTG 617
Db      124 LeuThrMetGlyMetAlaSerMetTyraArgGlySerAlaLeuAlaIleArgVal 143
Qy      618 CTGGTGTGGTGTGTGTCAGTACCGGCTGAGAAATATTGGTTCACCAAGGATCG 677
Db      144 ValValValThrIleGlnTyraArgLeuGlyValIleuGlyPhePheSerThrIleArgGln 163
Qy      678 CATGCTCCGGGAGACTGGGCTTCAAGACGAGAGTGGCTGTGTCTGCTGGGTCGAGAG 737
Db      164 HbAlaThrGlyAsnHbGlyTyraLeuArgGlnValAlaLeuArgTrrValGlnTyra 183
Qy      738 AACATGAGTTCTTTCGTTGGGAGCCCAAGCTGTGACCATCTTTGGCGAGTCCGCGGA 797
Db      184 AsnIleAlaHbPheGlyGlyAsnProGlyArgValThrIlePheGlyGlySerAlaGly 203
Qy      798 GCGATAGTGTTCATGCTTATCTGTCTGCTCCATGGCCAAAGCTTATTCACAAAGCC 857
Db      204 GlyThrSerValSerSerHbValIleuSerProMetSerGlnGlyLeuPheHbGlyAla 223
Qy      858 ATCATGAGAGAGTGGGCTGACATCCCTTACCTGAGAGCCCATGATTAAGAAGAGT 917
Db      224 IleMetGlySerLeuValAlaLeuLeuProGlyLeuIleThrSerSerGlnValVal 243
Qy      918 GAGGACSTGACGCTGCTGACATTTCTGTGTGAACAATGCTGACACTTGAAGCCCTG 977
Db      244 Ser-----ThrValValAlaAsnLeuSerArgCyGlyGlnValAspSerGlnThrLeu 261
Qy      978 CTGAGGTGCTGAGGACAAACCCCTCAAGAGCTGCTGACCTGACGCCAAAGAAACAAAG 1037
Db      262 ValArgCyValLeuArgAlaIleSerGlnGlnMetLeuAlaIleThrGlnValPheMet 281
Qy      1038 TCTTTCACSTGAGTGTGATGCTGCTTCTTCCATAGACCTCTGATGATTAATGCT 1097
Db      282 LeuIleProGlyValValAlaArgGlyValPheLeuProArgHbPheGlnGlyLeuLeuAla 301
Qy      1098 CAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATACCAAGAGTGGCTTC 1157
Db      302 LeuAlaAerPheGlnProValProSerIleIleGlyIleAsnAsnArgGlnTyraGlyTrr 321
Qy      1158 CTGCTGCTCT-----ATGAGAGAGCT 1178
Db      322 IleIleProGlyLeuLeuLeuAlaIleAerProGlnGlnGlyArgAspArgGlnAlaMet 341
Qy      1179 CTTGAGTCTCAAGTGGCTCCAAAGTCCCTTGGCTCCATCTGATACAAACATCTCG 1238
Db      342 ArgGlnIleMetHbGlnAlaThrTyraGlnLeuMetLeu----- 354
Qy      1239 CACATCCGCGCTCAGTATTTGACSTTGTGCTAATGATATCTTCATGACAAAGACTCC 1298
Db      355 -----ProProHbAlaLeuGlnTyraSerLeuMetAerGlnTyraMetGlySerAenIleAer 372
Qy      1299 CTGACTGAATTCGAGACAGTCTTTCGATCTTGGATGTGTTCTTTGTGGCTCT 1358
Db      373 ProGlyHbIleMetAlaGlnPheGlnGlnMetAlaAerAlaMetPheValMetPro 392
Qy      1359 GCACTGATCAAGCTCGATATACAGAGATGCTGTGATCTGCTTACTCTTATAGTTT 1418
Db      393 AlaLeuArgValAlaHbIleGlnTyra--SerHbAlaProThrTyraPheTyraGlnPhe 411
Qy      1419 CGGACCGGCTCAGTGTGTTGAAGACGAAAGCCGCTTTGTCAAGCCGACCAAGCT 1478

```

```

Db      412 GlnHbArgProSerPheThrTyraSerLeuArgProProHbAlaArgAlaPheHbGly 431
Qy      1479 GATGAAGTCCGCTTGTGTGCTGTGCTGCTGCTGCTGAGAGGAGACATTTGTTGCA 1538
Db      432 ArgGlnValValPheValPheArgSerHbIleuPheGlySerTyraValProIleu----- 449
Qy      1539 GGAGCCAGCGGAGAGAGAACTTACTGAGCCCGAAGATGATGAATATCTGGCTACTT 1598
Db      450 -----ThrGlnGlnGlnLeuLeuSerAlaGlyArgValMetTyraTrrIleAlaPhe 467
Qy      1599 GCTGACCCGAGGATCTTATGGAACAGCTGTCTGTGGCCAGCTTAAATATCTGACT 1658
Db      468 AlaArgAsnArgAsnProAsnGlyGlyLeuAlaHbIleTrrProLeuPheAerLeuAer 487
Qy      1659 GAGCAGTACCTCCACTGACTTGAACATGAGCCCTCGACAGAGCTCAAGACCGCG 1718
Db      488 GlnArgTyraLeuGlnLeuAenMetGlnProAlaValGlyGlnAlaLeuAlaArgArg 507
Qy      1719 GTGATTTTGGACGACGACCATCCCC 1745
Db      508 LeuGlnPheTrrThrHbIleThrLeuPro 516

RESULT 3
JC5408
carboxylesterase (EC 3.1.1.1) - human
C.Species: Homo sapiens (man)
C.Date: 10-Sep-1998 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C.Accession: JC5408
R.Schwer, H.; Langmann, T.; Daig, R.; Becker, A.; Aelalandid, C.; Schmitz, G.
Biochem. Biophys. Res. Commun. 233, 117-120, 1997
A.Title: Molecular cloning and characterization of a novel putative carboxylesterase, pre
A.Reference number: JC5408; MUID:97289502; PMID:9144407
A.Accession: JC5408
A.Molecule type: mRNA
A.Residues: 1-559 <SCH>
A.Cross-references: UNIPROT:000748; GB:Y09616; NID:92058317; PID:CAA70831.1; PID:9205831
A.Experimental source: intestine
C.Comment: This enzyme hydrolyzes many xenobiotics, such as carboxyl esters, thioesters &
C.Genetics:
A.Gene: GDB:CEB2; tCE; CE2
A.Cross-references: GDB:9959011
C.Superfamily: cholinesterase; cholinesterase homology
C.Keywords: carboxylic ester hydrolase; glycoprotein
F.58-544/Domain: cholinesterase homology <CH>
F.115-95,123-280,291-428/Disulfide bonds: #status predicted
F.111,276/Binding site: carbohydrate (Asn) (covalent) #status predicted
F.128,457/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,216-85 Length: 559
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 31.77% Indels: 24
DB: 1 Gaps: 8

US-10-023-515-1 (1-2158) x JC5408 (1-559)
Qy      186 GGGCTTGTGCTGAAGGCGCACAGAGAACAGGCTGGATGATTCAGGCGCAAGCA 245
Db      26 GlnGlnAerPheSerAlaSerProIleArgThrThrHbIleThrGlnValLeuGlySerLeu 45
Qy      246 GTCACTGTGCTGGAAGCCCTGTGCTGTAACGTGTCTTGGAGTCCCTTGTGCTGCT 305
Db      46 ValHbValTyraGlyAlaAsnAlaGlyValGlnThrPheLeuGlyIleProPheAlaTyra 65
Qy      306 CCGGCTGGGATCCCTGCTGATTTACGAACCGGAGCTGATGCGCTGGGATATCTTG 365
Db      66 ProProGlyGlyProLeuArgPheAlaProProGlnProGlnTyraSerGlyVal 85
Qy      366 CGAAGGCAAGCTCTTACCTTAATTTGTGCTCCAGAAC-----TCAGAG 410
Db      86 ArgAerGlyThrThrHbIleProAlaMetCyValLeuGlnAerLeuThrAlaValGlySerGln 505

```

```

QY 411 TGGCTGCTCTTATGATCAACATGCTCAAGTGCATTAACCCGAAATTGCGAGTGTCAAG 470
    |||
Db 106 PheLeu-----SerGlnPheAbnMetThrPheProSerAspSerMetSerGln 121
QY 471 GACTGCTCTTACCTGAACATCTATGCGCTGCCACCGCGATACAGGCTCCAGACTCC 530
    |||
Db 122 AspCysLeuTyrLeuSerIleTyrThrProLanHisSerHisGluGlySerAbnLeuPro 141
QY 531 GCTTGGTGGTGGTCCAGAGAGTGGCTTCAAGACTGCTCAAGCTTCATCTTGTATGG 590
    |||
Db 142 ValMetValIlePheHisGlyValAlaLeuValPheGlyMetAlaSerLeuTyrAspGly 161
QY 591 TCCGCTGCTGCTGCTGATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
    |||
Db 162 SerMetLeuAlaIleAlaLeuGluAbnValValValIleIleGlnTyrAbnLeuGlyVal 181
QY 651 TTTGGTTTCTTCAACACATGGAGATCAGATGCTCCGGGAACTGGGCTTCAAGACAG 710
    |||
Db 182 LeuGlyPhePheSerThrGlyAspLysHisIleAlaThrGlyAbnTyrGlyLeuAbnGln 201
QY 711 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
    |||
Db 202 ValAlaAlaLeuAbnGlyPheValGlnAbnHisIleAlaPheGlyGlyAbnProAspArg 221
QY 771 GTGACCATCTTTGGCGAGTCCGGCGAGCCATAGTCTTCTAGCTTCTTACTGCTGCC 830
    |||
Db 222 ValThrIlePheGlyGlnSerIleAlaGlyIleThrSerAlaSerIleValAlaSerPro 241
QY 831 ATGGCCCAAGGCTTATTCACAAAGCCATGATGAGAGTGGGCTGCCATCATCTTAC 890
    |||
Db 242 IISerGlnIleLeuPheHisGlyAlaIleMetGlnSerIleValAlaLeuLeuProGly 261
QY 891 CTGAGAGCCCATGATTATGAGAAAGTGAAGACCTG-----CAGTGGTGAACATTTT 944
    |||
Db 262 LeuIleAla-----SerSerAlaAspValIleSerThrValAlaAbnLeu 277
QY 945 TGTGATACAAATGCCGACACTTGAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
    |||
Db 278 SerAlaCysAbnGlnValAbnSerGlnAlaLeuValGlyCysLeuAbnGlyLysSerLys 297
QY 1005 AAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064
    |||
Db 298 GlnGlnIleLeuAlaIleAbnLysProPheLysMetIleProGlyValAlaAbnGlyVal 317
QY 1065 TTTCTTTCTTAATGAGCTCTAATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1124
    |||
Db 318 PheLeuProArgHisProGlnIleLeuAlaSerAlaAspPheGlnProValProSer 337
QY 1125 ATCATGAGATCAATTAACAGAGTGGCTTCTGCTGCTGCT-----ATGAAG 1172
    |||
Db 338 ILeValGlyValAbnAbnAbnGlnPheGlyTyrLeuIleProLysValMetArgIleTyr 357
QY 1173 GAGGCTCTGAGATCTCAGTGGCTCAACAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1232
    |||
Db 358 AspThrGlnLysGlnMetAspArgGlnLaseGlnAlaIleAlaLeuGlnLysMetLeuThr 377
QY 1233 ATCTGCAACATCCGCTCAGATTTGACCTTGGGCTTAATGAATCTTCCATGACAG 1292
    |||
Db 378 LeuLeuMetLeuProThrPheGlyAspLeuLeuArgGlnGlyIleGlyAspAbn 397
QY 1293 CACTGCTGATGAATCCGAGACAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1352
    |||
Db 398 GlyAspProGlnThrLeuGlnAlaGlnPheGlnIleMetIleAlaAspSerMetPheVal 417
QY 1353 GTCCCTGACATGATCAGCTGCTGATATCAGACAGATGCTGCTGCTGCTGCTGCTGCTGCT 1412
    |||
Db 418 IleProAlaLeuGlnValAlaHisPhe---GlnCysSerArgAlaProValIleTyrPheTyr 436
QY 1413 GAGTTTGGCAGACCGGCTCAGTGGCTTGAAGACAGACAGCGGCTTTGTCAAGCGGAC 1472
    |||
Db 437 GluPheGlnHisGlnProSerTyrPheLysAbnIleArgProProHisMetLysValAsp 456

```

```

QY 1473 CAGCTGATGAATCGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1532
    |||
Db 457 HisGlyAspGlnLeuProPheValPhe---ArgSerPhePheGlyValAbnTyrIleLys 475
QY 1533 TTGGAAGAGCCAGGAGAGAGAGAAATTACTGACCCGGAAGATGATAAATTACTGGCT 1592
    |||
Db 476 Phe-----ThrGlnIleGlnIleGlnIleLeuSerArgLysMetMetLysTyrTrpAla 492
QY 1593 ACCTTTCTGGAACCGGGAATCTTAATGGAACGACTGCTGCTGCTGCTGCTGCTGCTGCT 1652
    |||
Db 493 AsnPheAlaArgAbnGlnAbnProAbnHisGlnGlyLeuGlyLeuProHisTyrProLeuPheAsp 512
QY 1653 CTGACTGAGCACTACTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1712
    |||
Db 513 GlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 532
QY 1713 CCGCGGTGATTTTGGACAGCAACATCCCC 1745
    |||
Db 533 HisArgLeuGlnPheThrPheLysValAlaLeuPro 543

```

RESULT 4

S47655

carboxylesterase (EC 3.1.1.1) precursor - golden hamster

C.Species: Mesocricetus auratus (golden hamster)

C.Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C.Accession: S47655

R.Sone, T.; Isode, M.; Takabatake, E.; Wang, C.Y.

Biochim. Biophys. Acta 1207, 138-142, 1994

A.Title: Cloning and sequence analysis of a hamster liver cDNA encoding a novel putative

A.Reference number: S47655; PMID:94318665; PMID:8043605

A.Accession: S47655

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-561 <SON>

A.Cross-references: UNIPROT:Q64419; EMBL:D28566; NID:9531238; PIDD:BA05913.1; PIDD:953123

C.Superfamily: cholinesterase; cholinesterase homology

C.Keywords: carboxylic ester hydrolase

P.558-546/Domain: cholinesterase homology <CHI>

F.227,459/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.:	2.8e-83	Length:	561
Score:	1199.00	Matches:	252
Percent Similarity:	61.16%	Conservative:	74
Best local Similarity:	47.28%	Mismatches:	181
Query Match:	30.97%	Indels:	26
DB:	2	Gaps:	7

US-10-023-515-1 (1-2158) x S47655 (1-561)

```

QY 186 GGGCTTCTGCTGAAGGCGCACAGAGAACACAGGCTGGATGATTCAGGCGAAGCA 245
    |||
Db 26 GlyGlnAspSerValSerProLanHisSerGlnValAlaArgGlyLysLeu 45
QY 246 GTCACTGCTGCGAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
    |||
Db 46 ValTyrValLysGlnGlyValAlaThrGlyValTyrAlaPheLeuGlyIleProPheAlaLys 65
QY 306 CCGCGGTGATTTTGGACAGCAACATCCCC 1745
    |||
Db 66 ProProValGlyProLeuAbnPheAlaProProGlnProProGlnProProGlnProPro 85
QY 366 CGAAGACCACTCTTACCTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 425
    |||
Db 86 ArgAspGlyThrSerGlnProAlaMetCysLeuGlnThrAspPheMetArgProGlnIle 105
QY 426 CAACACATGCTCAAGCTGATTAACCGAAATTCGAGTGTCAAGAGTGTGCTTACTGCT 485
    |||
Db 106 SerLysGlnArgLysIleIleLeuProThrIleSerMetSerGlnAspCysLeuTyrLeu 125
QY 486 AACATCTATGCGCTGCGCCAGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 545
    |||
Db 126 AsnIleTyrThrProLanHisGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 145

```

Oy	546	CCAGAGGTCGCTTCAAACTGCGCTCAACCTCCATCTTGAATGGGTCGCGCTGGCTGCC	605
Db	146	HtGtIyGtAlaLeuValMetGlyMerLAsSerMetAenAapGlySerLeuLeuAlaLa	165
Oy	606	TATGAGACGTCGTCGTGTGGTGGTCCAGTAACCGGCTAGGAAATATTTGGTTCTTCAAC	665
Db	166	ThrtIuAapIleValIleValSerIleGlnTyrAgleuGlyIleLeuGlyPhePheSer	185
Oy	666	ACATGGATCAGCATGCTCCGGGGAACTGGGCTTCAAGGACAGAGTGGCTGTCTGCC	725
Db	186	ThrtIyAapGtIuHtIleAlaIaArgGlyAantTprGlyTyrLeuAapGlnValAlaIleuHtS	205
Oy	726	TGGGTCGCAAGAAACATGAGTCTTCCGATGGGGACCCCACTGTGACCATCTTGGC	785
Db	206	TtPValGtngInAenHtLeAlaSerPheGlyGlyAenProGlyGlnValThrIlePheGly	225
Oy	786	GAGTCGGGGGAGGCATAGTGTTCATGTCTTACTGTCTCCATGGCCAAAGGCTTA	845
Db	226	ValSerAlaGlyGlyThrSerValSerSerLeuValValSerProMetSerLysGlyLeu	245
Oy	846	TTCCACAAAGCCATCATGAGAGATGGGGTGGCCATCATCCCTTACCTGAGGCCCATGAT	905
Db	246	PheHtSerGlyAlaIleMetGlnSerGlyValAlaLeuLeuProAapLeu---IleSerAap	264
Oy	906	TATGAGAAAGATGAGAACCTGACGGTGGTGCACATTTCTGTGTAAACATGCTCGAAC	965
Db	265	ThtProGluAlaValIlyTyrThrProValAlaIaenGlnSerGlyCyGluAlaIlyAapS	284
Oy	966	TCTGAGGCGCTCGTAGGTCGCTGAGGACAAACCCCTCAAGAGTCTCAACCTCAGC	1025
Db	285	SerGtIuAlaLeuValHtSCyALeuAArgIuYbThrtGtIuAlaGlnIleLeuAlaIleAen	304
Oy	1026	CAGAAAACAAAGCTTTTCACTCGAGTGGTGTAGTGGCTTTCTTCCATAGAGCTCTA	1085
Db	305	GlnValPheIleMetThrProGlyValValAapGlyIlePheLeuProAapHtIleProGln	324
Oy	1086	GATCTATTGTCTCAGAAAGCATTTAAAGCAATCCCTTCATCATCGAGTCAATAAACCA	1145
Db	325	GluLeuLeuAlaSerValAapPheHtIleProValProSerIleIleGlyValAapSerAap	344
Oy	1146	GAGTGTGGCTTCTGCTGCTCATG-----AAGAGGCT	1178
Db	345	GluCySbGlyTtPrgIyValProLeuPheMetGlyLeuAapHtIleValIleIyAaenHtIleThr	364
Oy	1179	CCTGAGATCCTCAGTGGCTCCAAAGTCCCTTGCCCTCATCTGATCAAAACATCCTG	1238
Db	365	ArgGtIuHtIleuProAlaPheLeuIySerAArgAlaGhtIuHtIleMet-----	380
Oy	1239	CACATCCCGCTCAGTATTTGACCTTGAGCTAATGAATCTTCATGACAAACACTCC	1298
Db	381	---LeuProProGluCySerAapLeuLeuMetGlnIuTyrMetGlyAapValGluAap	399
Oy	1299	CTGACGTAAATCCAGACAGTCTTCTGCACTTCTGAGATGATGTTCTTGTGGTCCCT	1358
Db	400	ProGlnThrLeuGlnAlaGlnPheAArgGlyLeuMetLysAapPheMetPheValIlePro	419
Oy	1359	GCATGATCAGAGCTCGATATACAGAGATGCTGTGACCTGTCATCTTATGAAGTT	1418
Db	420	AlaLeuIyValAlaIlyPheGlnAArg---SerHtIleAProValIlyPheTyrGluPhe	438
Oy	1419	CGGACACGGCTCAGTCTT-----GAAACACAGAACCGGCTTTGTCAAGCCGAC	1472
Db	439	GlnHtIleGlnSerSerPheHtIleYAsnIyAapAlaArgProSerHtIleValAlArgAlaAap	458
Oy	1473	CACGCTATGAAGTCCGCTTGTGTGTGGTGGTGCCTT-----CTGAAGGGGACATT	1526
Db	459	HtGtIyAapHtIleValAlaPheValPheGlySerAapPheTtPrgIyLeuIyHtIleAapLeu	478
Oy	1527	GTTATGTTCAAGAGCAACGAGAGAGAAAGTATCTGAGCCGGAAGATGATGAATAC	1586
Db	479	-----ThrtGlnGlnIuIyIySerLeuAanIyAArgMetIleIyTyr	492

```

OY      1587  TGGGCTACCTTGGCTCGAACC GGGAATCCTAAATGGGAAGACACCTGTCTCTGGCAGCT 1646
Db      493  TTPAlaasnpHnAlaIrgHlscgIyAsnpProAsSercIuclYLeuProIyrtITpProIu 512
OY      1647  TATAATCTGACTGAGCAGCAGTACTCCAGCTGACTGACTGAAACATGACCTCGAGCAGAGACTC 1706
Db      513  LeuValhIsAspAlaSpclnIrtYLeuIySleuAspIleclnProAlaValGIyAgAlaLeu 532
OY      1707  AAAGAACCGCGGGGTGATTTTGGACCAGACCATCCCC 1745
Db      533  LySserArgLySLeuHIspeHtPrItnLySleLeuPro 545

RESULT 5
S34607
carboxylesterase (BC 3.1.1.1) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S34607
R:Aida, K.; Moore, R.; Negishi, M.
B:Biochim. Biophys. Acta 1174, 72-74, 1993
A:Title: Cloning and nucleotide sequence of a novel, male-predominant carboxylesterase in
A:Reference number: S34607; MUID:93326638; PMID:7916639
A:Accession: S34607
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <AID>
A:Cross-references: UNIPROT:063880; GB:S64130; NID:G404388; PIDN:AAE27606.1; PID:G404389
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
C:46-536/Domain: cholinesterase homology <CH>
C:215,443/Active site: Ser, His #status predicted

```

Alignment Scores:			
Record No.:	2,526-82	Length:	554
Score:	1186.50	Matches:	256
Percent Similarity:	61.16%	Conservative:	81
Best Local Similarity:	46.46%	Mismatches:	183
Query Match:	30.64%	Indels:	31
DB:	1	Gaps:	10

US-10-023-515-1 (1-2158) x S34607 (1-554)

QY	132	TGCTTTTTCGTGATTTCTCCAGCCCTGTGGGACACAGACAGTGGGAAAAATGGGCTT	191
DB <td>3 <td>CysLeuLeuLeuIleIlePheProThrThrVal-----IleGlyPro</td> <td>15</td> </td>	3 <td>CysLeuLeuLeuIleIlePheProThrThrVal-----IleGlyPro</td> <td>15</td>	CysLeuLeuLeuIleIlePheProThrThrVal-----IleGlyPro	15
QY <td>192 <td>TCTGTGTAAGGGCCACAGAGAACACACAGCTGGATGTGATTACAGGCGAACAGTCACT</td> <td>251</td> </td>	192 <td>TCTGTGTAAGGGCCACAGAGAACACACAGCTGGATGTGATTACAGGCGAACAGTCACT</td> <td>251</td>	TCTGTGTAAGGGCCACAGAGAACACACAGCTGGATGTGATTACAGGCGAACAGTCACT	251
DB <td>16 <td>LysValThrGlnProGlnValAlaThrProLeuGlyArgValArgGlnValGly</td> <td>35</td> </td>	16 <td>LysValThrGlnProGlnValAlaThrProLeuGlyArgValArgGlnValGly</td> <td>35</td>	LysValThrGlnProGlnValAlaThrProLeuGlyArgValArgGlnValGly	35
QY <td>252 <td>GTGCTGGGAAGCCCTGTGCTGTGAACGTTCTCTGGAGTCCCTTGTGCTGCTCCCG</td> <td>311</td> </td>	252 <td>GTGCTGGGAAGCCCTGTGCTGTGAACGTTCTCTGGAGTCCCTTGTGCTGCTCCCG</td> <td>311</td>	GTGCTGGGAAGCCCTGTGCTGTGAACGTTCTCTGGAGTCCCTTGTGCTGCTCCCG	311
DB <td>36 <td>ValLysAspThrAspArgMetValAsnValPheLeuGlyIleProPheAlaGlnAlaPro</td> <td>55</td> </td>	36 <td>ValLysAspThrAspArgMetValAsnValPheLeuGlyIleProPheAlaGlnAlaPro</td> <td>55</td>	ValLysAspThrAspArgMetValAsnValPheLeuGlyIleProPheAlaGlnAlaPro	55
QY <td>312 <td>CTGGATCCCTGGCATTTTACAGAACCCGACGCTGCATCCGCTGGATTACTTGGAGAA</td> <td>371</td> </td>	312 <td>CTGGATCCCTGGCATTTTACAGAACCCGACGCTGCATCCGCTGGATTACTTGGAGAA</td> <td>371</td>	CTGGATCCCTGGCATTTTACAGAACCCGACGCTGCATCCGCTGGATTACTTGGAGAA	371
DB <td>56 <td>LeuGlyProLeuArgPheSerAlaProLeuProProGlnProGlnProGlnValAlaArgAsp</td> <td>75</td> </td>	56 <td>LeuGlyProLeuArgPheSerAlaProLeuProProGlnProGlnProGlnValAlaArgAsp</td> <td>75</td>	LeuGlyProLeuArgPheSerAlaProLeuProProGlnProGlnProGlnValAlaArgAsp	75
QY <td>372 <td>GCCACCTCCATCCATTAATTGTGCTCCAGAACTAGACAGTGGCTGCTTAAATCAACAC</td> <td>431</td> </td>	372 <td>GCCACCTCCATCCATTAATTGTGCTCCAGAACTAGACAGTGGCTGCTTAAATCAACAC</td> <td>431</td>	GCCACCTCCATCCATTAATTGTGCTCCAGAACTAGACAGTGGCTGCTTAAATCAACAC	431
DB <td>76 <td>AlaSerIleAsnProMetCysLeuGlnAspValGlyArgMetSerAsnSerArgPhe</td> <td>95</td> </td>	76 <td>AlaSerIleAsnProMetCysLeuGlnAspValGlyArgMetSerAsnSerArgPhe</td> <td>95</td>	AlaSerIleAsnProMetCysLeuGlnAspValGlyArgMetSerAsnSerArgPhe	95
QY <td>432 <td>ATGCTCAGAGTGCATTACCCGAAATTGCGAGTGTGAGAAAGTCCGCTACCTGAACATC</td> <td>491</td> </td>	432 <td>ATGCTCAGAGTGCATTACCCGAAATTGCGAGTGTGAGAAAGTCCGCTACCTGAACATC</td> <td>491</td>	ATGCTCAGAGTGCATTACCCGAAATTGCGAGTGTGAGAAAGTCCGCTACCTGAACATC	491
DB <td>96 <td>ThrLeuAsnGlnLysMetLysIlePheProIleSerGlnLysCysLeuThrLeuAsnIle</td> <td>115</td> </td>	96 <td>ThrLeuAsnGlnLysMetLysIlePheProIleSerGlnLysCysLeuThrLeuAsnIle</td> <td>115</td>	ThrLeuAsnGlnLysMetLysIlePheProIleSerGlnLysCysLeuThrLeuAsnIle	115
QY <td>492 <td>TATGCGCTGCCACGCGCATACAGGCTCCAGGCTCCAGGCTCCGCTTGTGTGATTTCCAGAA</td> <td>551</td> </td>	492 <td>TATGCGCTGCCACGCGCATACAGGCTCCAGGCTCCAGGCTCCGCTTGTGTGATTTCCAGAA</td> <td>551</td>	TATGCGCTGCCACGCGCATACAGGCTCCAGGCTCCAGGCTCCGCTTGTGTGATTTCCAGAA	551
DB <td>116 <td>TySerProThrGlnIleThrAlaGlyAspLysArgProValMetValTrpIleAsnGly</td> <td>135</td> </td>	116 <td>TySerProThrGlnIleThrAlaGlyAspLysArgProValMetValTrpIleAsnGly</td> <td>135</td>	TySerProThrGlnIleThrAlaGlyAspLysArgProValMetValTrpIleAsnGly	135
QY <td>552 <td>GGTGCTTCAAGACAGTGGCTCAGGCTCATCTTGTGAATGGGTGGGCTGGCTGGCTAAGAG</td> <td>611</td> </td>	552 <td>GGTGCTTCAAGACAGTGGCTCAGGCTCATCTTGTGAATGGGTGGGCTGGCTGGCTAAGAG</td> <td>611</td>	GGTGCTTCAAGACAGTGGCTCAGGCTCATCTTGTGAATGGGTGGGCTGGCTGGCTAAGAG	611
DB <td>136 <td>GlySerLeuAspValGlySerSerThrSerAspGlySerAlaLeuAlaValatGlyGly</td> <td>155</td> </td>	136 <td>GlySerLeuAspValGlySerSerThrSerAspGlySerAlaLeuAlaValatGlyGly</td> <td>155</td>	GlySerLeuAspValGlySerSerThrSerAspGlySerAlaLeuAlaValatGlyGly	155

```

QY 612 GACGTCGTGGTGGTCCAGTACCGGGCTAGGAATTTGGTTCTTCAACGATGG 671
DB 156 AapValValValValThValValGlnTrpAglLeuGlyLeuPheGlyPheLeuSerThr 175
QY 672 GATCAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGTCTGCTCCGGGTC 731
DB 176 AapLysH1AmeCProGlyAasnArgGlyPheLeuAapValValAlaAlaLeuAglTrpVal 195
QY 732 CAGAAGAACATGAGTCTTGGTGGGAGCCCAAGCTGTGACCATTTTGGCGAAGTCC 791
DB 196 GlnGlyAanIleAaProPheGlyGlyAapProAanCyValIlePheGlyAasnSer 215
QY 792 GCGGAGCCATTAAGTGTTCATGCTTACTGTCTCCAGGCGCAAGGCTTATTCAC 851
DB 216 AlaGlyGlyIleIleValSerSerLeuLeuLeuSerProMetSerAlaGlyLeuPheH1s 235
QY 852 AAAGCCATCATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGAGCCCATGATTATGAG 911
DB 236 ArgAlaIleSerGlnSerGlyValValIleSerLysIleLeuGln-----AapLeuAan 253
QY 912 AAGAGTGAAGACCTGAGGTGGTGCACAT-----TTCTGTGGTAAACAATGGCTCAGAC 965
DB 254 AlaTrpSerGlnAlaGlnAasnPheAlaAasnValAlaCyGly---SerAlaSerPro 272
QY 966 TCTGAGGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGCTGTACCTCAGC 1025
DB 273 AlaGln---LeuValGlnCyLeuLeuGlnLysGlnGlyLysAapLeuIleThrLys 291
QY 1026 CAGAAAACAAGTCTTCACTGAGTGGTGGTGGCTTCTTCTTCAATGAGCTCTTA 1085
DB 292 AasnValAanIleSerLysTrp-----ValAasnAapSerPhePheProGlnArgProGln 309
QY 1086 GATCTATTGCTCTCAAGAAAGCATTTAAAGCAATCTTCCATCATCCGAGTCAATAACAC 1145
DB 310 LysLeuLeuAlaAasnLysGlnPheProThrValProLysLeuLeuGlyValIleThrAanH1s 329
QY 1146 GAGTGTGGTCTCCGCTG-----CCTTGAAGAG---GCT 1178
DB 330 GlnPheGlyTrpLeuLeuLeuLysPheTrpAanIleLeuAapLysMetGlnH1AlysLeuSer 349
QY 1179 CCTGAGATCTCATGCTCCCAAGAGCTCTTCCCTCCATCATGATCAAAACATCCG 1238
DB 350 GlnGlnAapLeuLeuGlnAasnSerArgProLeuLeuAlaH1s-----Met 364
QY 1239 CACATCCGCGCTCAGATTTGCACTTGGCTGAATGAATATTCTTCATGACAAAGCATCC 1298
DB 365 GlnLeuProProGlnIleMetProThrValIleAapGlyLysLeuAapAanGlySerAap 384
QY 1299 CTGACTGAATCCGAGACAGTCTTCTGACTGCTTGGAGTGGTCTTGGTGGTCCCT 1358
DB 385 GlnSerAlaThrArgTrpAlaLeuGlnGlnLeuLeuGlyAapIleThrLeuValIlePro 404
QY 1359 GCACGTATCAACGCTCGATTCACAGAGATGCTGTGACCTGTCTACTTATGAGTTT 1418
DB 405 ThrLeuIlePheSerLysTrpLeuGlnAapAlaGlyCyProValPheLeuTrpGlnPhe 424
QY 1419 CGGACCGGCGCTCAGTGTCTTGAAGACAGAACCGGCTTTTGCACAAAGCGACACGCT 1478
DB 425 GlnH1AthrProSerSerPheAlaLysPheLysProAlaTrpValLysAlaAapH1Ser 444
QY 1479 GATGAAGTCCGCTTGTGTGGTGGTGGCTTCTG-----AAGGGGACATGTTATG 1532
DB 445 SerGlnAanAlaPheValPheGlyGlyProPheLeuThrAapGlnUserSerLeuLeuAla 464
QY 1533 TTGCAAGAGCCAGGAGGAGAGAGATTAATCTGAGCCGGAAGATGAATAATCTGGGCT 1592
DB 465 PheProGlnAaThrGlnGlnGlnLysGlnLysLeuSerLeuThrMetAlaGlnTrpSer 484
QY 1593 ACCTTGTCTCAACCGGGAATCTTAATGGGAAACAGCTGTCTGTGAGCCAGCTTAAAT 1652
DB 485 GlnPheAlaArgThrGlnLysAasnProAanGlyLysGlyLeuProProTrpProGlnLeuAan 504

```

```

QY 1653 CTGACTGACAGTAACTCCAGCTGAGCTTGAACATGAGCTCCGACAGAGACTCAAGAA 1712
DB 505 GlnLeuGlnGlnTrpLeuGlnIleGlyLeuGlnProAaGlnTrpGlyValLysLeuLys 524
QY 1713 CCGCGGTGAGATTTTGGACCAAGCAACATCCCT 1745
DB 525 GlyAglLeuGlnPheTrpThrGlnTrpLeuPro 535

RESULT 6
A39060
carboxylesterase (EC 3.1.1.1) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 09-Jul-2004
C/Accession: A39060
R/Ovnic, M.; Tepperman, K.; Medda, S.; Elliott, R.W.; Stephenson, D.A.; Grant, S.G.; Gane
Genomics 9, 344-354, 1991
A/Title: Characterization of a murine cDNA encoding a member of the carboxylesterase mult
A/Reference number: A39060; MUID:91169540; PMID:1840565
A/Accession: A39060
A/status: preliminary
A/molecule type: mRNA
A/residues: 1-554 <OVN>
A/cross-references: UNIPROT:P23953; GB:M57960; NID:9192853; PIDN:AAA63297.1, PID:9192854
C/superfamily: cholinesterase; cholinesterase homology
C/keywords: carboxylic ester hydrolase
F:50-540/Domain: cholinesterase homology <CH>
F:221,455/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,57e-80 Length: 554
Score: 1163.00 Matches: 241
Percent Similarity: 59.74% Conservative: 84
Best Local Similarity: 44.30% Mismatches: 193
Query Match: 30.04% Indels: 26
DB: 2 Gaps: 7

US-10-023-515-1 (1-2158) x A39060 (1-554)
QY 141 CTGATTTCTCAAGCCCTGTTGGACACAGACAGTGGGAAAACTGGGCTTCTGTGAA 200
DB 11 LeuAlaValCyProIleLeuGlnH1s-----SerLeuLeu 22
QY 201 GGGCCACAGAGAACCAAGAGCTGGAGTATTCAGGCAAGATGATCATCTGCTGGGA 260
DB 23 ProProValValAapThrTrpGlnGlyLysValLeuGlyLysTrpLysLeuGlnGly 42
QY 261 AGCCCTGTGCTGTGAAGTGTCTCTGGAGTCCCTTGTCTGCTCCCGCTGGATCC 320
DB 43 PheGlnGlnProValAlaValPheLeuGlyValIlePheAlaLysProProLeuGlySer 62
QY 321 CTGCGATTTACGAACCGGAGCTGCATGCGCTGGATGATTAATCTCCGAAAGCCACCTCC 380
DB 63 LeuAapPheAlaProProGlnProAlaGlnProTrpSerPheValLysAanAlaThrSer 82
QY 381 TACCTTAATTTGTGCTCCAGAACTCAAGTGG-----CTGCTTTAGATCAACATG 434
DB 83 TrpProPheMetCySerGlnAapAlaGlyTrpAlaLysIleLeuSerAapMetPheSer 102
QY 435 CTCAAGGTGCATTAACCGAAATTCGAGTGTCAAGAACTGCTCTTACTTAACATTTAT 494
DB 103 ThrGlnLysGlnIleLeuProLeuLysIleSerGlnAapCyLysLysLeuAanIleTrp 122
QY 495 GCGCTGCCCAAGCCGATACAGCTCAAGCTCCCGCTTGTGGTGGTCCAGAGAGT 554
DB 123 SerProAlaAapLeuThrLysSerSerGlnLeuProAlaMetValTrpIleH1GlyGly 142
QY 555 GCGTTCAAGACTGGGTCAAGCTCCATCTTGAATGGTCCGCTGCTGCTATGAGAG 614
DB 143 GlyLeuValIleGlyGlyArgSerProLysAanGlyLeuAlaLeuSerAlaH1AglAan 162
QY 615 GTGCTGTTGTGTGTCACGATACCGGCTAGGAATATTGTTCTTACACACATGGAGT 674
DB 163 ValValValValThrIleGlnTrpAglLeuGlyIleTrpGlyLeuPheSerThrGlyAap 182

```

```
Oy 675 CAGCATGCTCCGGGAACTGGGCTTCAAGACAGAGTGGCTGTGTCCTGGTCCG 734
Db 183 GlnHisSerProGlyAsnThrPheAlaHisLeuAspGlnLeuAlaHisLeuArgTrpValGln 202
Oy 735 AAGAACAATGAGTTCTTGGGTGGGAGCCCAAGCTGTGACCATTTTGGCGAGTCCG 794
Db 203 AsnProAlaHisLeuPheGlyGlyLeuAsnProAspSerValThrIlePheGlyGlyLeuSer 222
Oy 795 GGAGCCATTAAGTGTCTTACTGTCTTACTGTCTTACTGTCTTACTGTCTTACTGTCTT 854
Db 223 GlyGlyIleSerValSerValLeuValLeuSerProLeuGlyValAspLeuPheHisArg 242
Oy 855 GCCATCATGAGAGGAGGGGGGCGCATCCCTTACCTGAGAGGCCCATGATATGAGAG 914
Db 243 AlaIleSerGlnSerGlyValValIleAsnThrAsnValGlyValLeuAsnIleGlnAla 262
Oy 915 AGTGAAGACCTGACAGTGGTTGCAATTTCTGTGTAAACAATGCGTCAAGACTGTAGGCC 974
Db 263 ValAsnGlnIleIleAlaThrLeuSerGlnCys-----AsnAspThrSerSerAlaAla 280
Oy 975 CTGCTGAGGTGCTGAGACAAACCTTCAGAGGCTGTGACCTCAGC----- 1025
Db 281 MetValGlnCysLeuArgGlnIleSerThrGlnSerGlnLeuGlnIleSerGlyValSer 300
Oy 1026 ---CAGAAACAAAGCTTTTCACTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1082
Db 301 ValGlnIleAsnIleSerLeuSerThrMetIleAspGlyValValLeuProValAspPro 320
Oy 1083 CTAGATCATTTGTCTCAGAAACATTTAAAGCATTTCTTCCATCATGAGAGTCAATAC 1142
Db 321 GlnGlnIleLeuAlaGlnIleSerPheAsnThrValProIleValGlyPheAsnIle 340
Oy 1143 CAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1190
Db 341 GlnGlnPheGlyTrpIleIleProMetLeuGlnAsnLeuLeuProGlnIleValSerMet 360
Oy 1191 AGTGCTTCAACAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1250
Db 361 AsnGlnIleThrAlaSerLeuLeuLeuArgArgPheHisSerGlnLeuAsnIleSerGln 380
Oy 1251 CAGTATTTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1310
Db 381 SerMetIleProAlaValIleGlnIleThrValLeuArgValAspAspProAlaValSer 400
Oy 1311 CGAAGACAGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1370
Db 401 SerGlnLeuIleLeuAspMetPheGlyAspIlePhePheGlyIleProAlaValLeuLeu 420
Oy 1371 GCTGATATTCACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1430
Db 421 SerArgSerLeuAspAspAlaGlyValSerThrIleMetIleGlnPheArgTrpArgPro 440
Oy 1431 CAGTGTCTTGAAGACACAGAGCGGCTTTTGTCAAGGCGACACGCTGATGAAGTCCGC 1490
Db 441 SerPheValSerAspValArgProGlnThrValGlnGlyAspHisGlyAspGlnIlePhe 460
Oy 1491 TTTGTGTGCGGTGCTGCTTCTGTAAGGGGACATTTGTTATGTTGAAGAGCCAGGAG 1550
Db 461 PheValPheGlyAlaProLeuLeuLys-----GlnGlyValSerGln 474
Oy 1551 GAGGAGAAAGTAACTGAGCGCGAGAGATGAATACTGGGCTACCTTTGCTGGAACCGGG 1610
Db 475 GlnGlnIleThrAsnLeuSerIleMetValMetLysPheTrpAlaAsnPheAlaArgAsnGly 494
Oy 1611 AATCTTAATGGAAGACGCTGTCTGTGCGAGCTTAACTGATGCTGAGCAGTACCTTC 1670
Db 495 AsnProAsnGlyGlnGlyLeuProHisTrpProGlnIleValAspGlnGlnIleValLeu 514
Oy 1671 CAGCTGACCTTGAAACATGAGCTCGGACAGAGACTCAAGAAACCGCGGGTGAATTTTGG 1730
Db 515 GlnIleGlyAlaThrThrGlnGlnAlaGlnArgLeuValGlnGlyValAlaIlePheTrp 534
```

```
Oy 1731 ACCAGACCAATC 1742
Db 535 ThrGlnLeuLeu 538

RESULT 7
S19307
carboxylesterase (EC 3.1.1.1) precursor - pig
N:Alternate names: proline-beta-naphthylamidase
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S19307; S23607
R:Matsumi, M.; Inoue, H.; Ichinose, M.; Teukada, S.; Miki, K.; Kurokawa, K.; Takahashi
FEBS Lett. 293, 37-41, 1991
A>Title: The nucleotide and deduced amino acid sequences of porcine liver proline-beta-n
A:Reference number: S19307; MUID:92070571; PMID:1959668
A:Accession: S19307
A:Molecule type: mRNA
A:Residues: 1-566 <MAT1>
A:Cross-references: UNIPROT:Q29550; EMBL:X63323; NID:g1930; PID:CAA44929.1; PID:g1931
A:Accession: S23607
A:Molecule type: protein
A:Residues: 19-40 <MAT2>
A>Note: 28-Lys and 33-Leu were also found
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-566/Product: carboxylesterase #status experimental <MAT>
F:51-552/Domain: cholinesterase homology <CHS>
F:80/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:222,467/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,72e-80 Length: 566
Score: 1162.50 Matches: 250
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 44.33% Mismatches: 202
Query Match: 30.02% Indels: 41
DB: 2 Gaps: 8

US-10-023-515-1 (1-2158) x S19307 (1-566)
Oy 129 TGTGCTTTTCTGCTGATTCCTCAGCCCTGTGGAGACAGACAGTGGGAAAACTGGG 188
Db 2 TrpLeuLeuProLeuValLeuThrSerLeuAlaSerSerAlaThrTrp-----AlaGly 19
Oy 189 CTTTGTCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGATGC 248
Db 20 GlnProAlaSerProProValValAspThrAlaGlnGlyArgValLeuGlyVal 39
Oy 249 ACTGTCGTGGAAAGCCCTGTGCTGTAACCGTTCCTCGAGGTCCCTTTGCTGCTGCC 308
Db 40 SerLeuGlnGlyLeuAlaGlnProValAlaValPheLeuGlyValProPheAlaValPro 59
Oy 309 CCGCTGGATCCCTCGATTTACGAACCCGAGAGCTGTGATCGCCCTGGATTAATCTGGA 368
Db 60 ProLeuGlnSerLeuAspArgPheAlaProProGlnProAlaGlnProTrpSerPheValLys 79
Oy 369 GAAGCACCTCTACCTTAATTGTGCTCCAGAAC-----TCA 407
Db 80 AsnThrThrSerIleProProMetCysCysGlnAspProValValGlnIleMetThrSer 99
Oy 408 CAGTGGCTGCTTATGATCAACACATGCTCAAGTGTCAAGTTCACCGAAATTCGAGTGC 467
Db 100 AspLeuPheThrAsnGlyValGlnArgLeuThrLeuGlnPhe-----Ser 114
Oy 468 GAAGACTCCCTTACTGTAATCATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
Db 115 GlnAspCysLeuIleValLeuAsnIleThrThrProAlaAspLeuThrIleValArgGlyValLeu 134
Oy 528 CCGGTCTTGTGTGCTGCTTCCAGAGAGTCCCTTCAAGACTGCTCAGCTCATCTTTGAT 587
Db 135 ProValMetValTrpIleHisGlyGlyValLeuValLeuGlyGlyAlaProMetIleArg 154
```


Db 138 TrpIleHisGlyGlyLeuIleIleGlyAlaSerProTyrSerGlyLeuAlaLeu 157
 QY 600 GTCGCTATGAGACGTCGTGGTGGTCCGCTGACGCTACCGCTGAAATATTTGGTTTC 659
 Db 158 SerIleHisGlyAlaSerValValValThrIleGlyTyrArgLeuGlyIleTrpIleLeu 177
 QY 660 TTCAACATGATGATCAGATGCTCCGGGAACTGGGCTTCAAGACAGACGAGCTGCT 719
 Db 178 PheSerThrGlyArgGluHisSerArgIleSerThrAlaHisIleLeuArgGluAla 197
 QY 720 CTGTCTCTGGTCCAGAGACATGAGTTCTTGGTGGGAAACCCGACGCTGTGTGACATC 779
 Db 198 LeuArgTrpValGlnAspAsnIleAlaAsnPheGlyGlyAsnProAspSerValThrIle 217
 QY 780 TTGGGCGATCGCGGGGAGCAATAGTTTCTACTCTTATATCTGTCCCATGGCCAA 839
 Db 218 PheGlyGlySerAlaGlyGlyValSerValSerAlaValLeuSerProLeuAla 237
 QY 840 GAGCTTATTCACAAAGCATGATGAGAGTGGGCTGGCCATCATCCCTTACCTGAGGCC 899
 Db 238 AsnLeuPheHisArgAlaIleSerGluSerGlyValIleLeuThrAsnLeuAspLys 257
 QY 900 CATGATTTATGAAAGATGAGACCTCGAGCTGTGACATTTCTGTGTAAACATGCG 959
 Db 258 LysAsnThrGlnAlaVal-----AlaGlnMetIleAlaThrLeuSerGlyCysAsnAsn 275
 QY 960 TCAGACCTGAGGCGCTGCTGAGCTGCTGAGGAGAAACCTCCAGAGCTGTGACC 1019
 Db 276 ThrSerSerAlaAlaMetValGlnCysLeuArgGlnLysTrpGluAlaGluLeuLeu 295
 QY 1020 CTCAGCCGAGAA-----ACAAGCTCTTCACTCGAGCTGTGATGCTGCTTTCTTCT 1073
 Db 296 LeuThrValLysLeuAspAsnThrSerMetSerThrValIleAspGlyValValLeuPro 315
 QY 1074 AATGAGCTCTAGATCTATTGTCTCAGAAACATTTAAAGCATTCCTTCATCATCGGA 1133
 Db 316 LysThrProGluGluIleLeuThrGluLysSerPheAsnThrValProTyrIleValGly 335
 QY 1134 GTCAATACACAGATGCTGCTCTGCTGCTATGAAAGAGGCTCCGATGATCCTCATG 1193
 Db 336 PheAsnLysGlnGluPheGlyTyrIleIleProThrMetMetGlyAsnLeuLeuSerGlu 355
 QY 1194 GGC---TCCAAACAGTCCCTGCGCTCCATCATGATACAAACATC-----CTGAC 1241
 Db 356 GlyArgMetAsnGluLysMetAlaSerSerPheLeuLysArgPheSerProAsnLeuAsn 375
 QY 1242 ATCCCGCTCAGTATTTGACCTTGTGCTATGATGATCTTCATGACAGACATCCCTG 1301
 Db 376 IleSerGluSerValIleProAlaIleIleGluLysTyrLeuArgGlyTyrAspAspPro 395
 QY 1302 ACTGAATCCAGACAGATCTTGTGACCTTGTGAGATGTTGTTTGTGGTCCCTGCA 1361
 Db 396 AlaLysLysLysGluLeuLeuLeuAspMetPheSerAspValPhePheGlyIleProAla 415
 QY 1362 CTGATCAGACGCTGATCAGAGATGCTGAGACCTGTCATCTTCTATGAGTTGG 1421
 Db 416 ValIleMetSerAspArgSerLeuArgAspAlaGlyAlaProThrTyrMetTyrGluPheGln 435
 QY 1422 CACCGGCTCAGTGTGTAAGACAGACGAGCGGCTTTTGTCAAGCCGACACGCTGAT 1481
 Db 436 TyrArgProSerPheValSerAspGlnArgProGlnThrValGlnGlyAspHisGlyAsp 455
 QY 1482 GAAGTCCGCTTGTGTTGCTGCTGCTCTTCTGAAAGGGGACATTTGTTATGTTGCAAGA 1541
 Db 456 GluIlePheSerValPheGlyTyrProPheLys-----GluGly 469
 QY 1542 GCCACGAGAGGAGAACTTATGAGCGGAGAAATGAAATACCTGGGCTACCTTGTCT 1601
 Db 470 AlaSerGluGluGluThrAsnLeuSerLysLeuValMetLysPheTrpAlaAsnPheAla 489
 QY 1602 CGAACCGGAAATCTTATGAGACAGACCTGTCTGTGGCCAGCTTATATCTGACGAG 1661
 Db 490 ArgAsnGlyAsnProAsnGlyGluGlyLeuProHisTrpProGluTyrAspGlnLysGlu 509

QY 1662 CAGTACTCCAGCTGACTTGAACATGAGCTCCGACAGACATCCGAGCGGCTG 1721
 Db 510 GlyTyrLeuGlnIleGlyAlaThrThrGlnGlnAlaGlnLysLeuGlyGluGluVal 529
 QY 1722 GATTTTGGACGACGACCATC 1742
 Db 530 AlaPheTrpThrGluLeu 536
 RESULT 9
 A1010
 A:Enzyme: EC 3.1.1.1 precursor, monocyte/macrophage (validated) - human
 N:Alternate names: carboxylesterase, hepatic; monocyte/macrophage serine esterase
 C:Species: Homo sapiens (man)
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
 R:Accession: A41010; UN0327; A47376; A49816; P80280; I61085; A48809; I57004
 J:Rat: J. S.; Shi, G.P.; Marx, E.A.; Chin, D.T.; Gerard, C.; Chapman, H.A.
 J:Bio: J. Chem. 266, 18832-18838, 1991
 A:Title: A serine esterase released by human alveolar macrophages is closely related to 1
 A:Reference number: A41010; MUID:92011649; PMID:1918003
 A:Accession: A41010
 A:Molecule type: mRNA
 A:Residues: 1-567 <MUN>
 A:Cross-references: UNIPROT:P23141; GB:M73499; NID:G179927; PID:AAA5649.1; PID:G179928
 A:Note: parts of this sequence, including the amino end of the mature protein, were confi
 R:Long, R.M.; Calabrese, M.R.; Martin, B.M.; Pohl, L.R.
 Life Sci. 48, PL43-PL49, 1991
 A:Title: Cloning and sequencing of a human liver carboxylesterase isoenzyme.
 A:Reference number: UN0327; MUID:91148424; PMID:1997784
 A:Accession: UN0327
 A:Molecule type: mRNA
 A:Residues: 61-567 <ION>
 A:Cross-references: GB:M55509; NID:G179929; PID:AAA5650.1; PID:G179930
 R:Shibata, F.; Takagi, Y.; Kikajima, M.; Kuroda, T.; Omura, T.
 Genomica 17, 76-82, 1993
 A:Title: Molecular cloning and characterization of a human carboxylesterase gene.
 A:Reference number: A47376; MUID:94010913; PMID:8406473
 A:Accession: A47376
 A:Structure: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA; DNA
 A:Residues: 1-3, 'PALV', 8-11, 'A', 13-567 <SHI>
 A:Cross-references: GB:D21088; NID:G455476; PID:BA04650.1; PID:G458470
 A:Note: sequence extracted from NCBI backbone (NCBI:137630) and corrected to correspond
 R:Zschocke, F.; Salmasi, A.; Kreize, H.; Buck, F.; Patzwesch, M.R.; Radzun, H.J.
 Blood 78, 506-512, 1991
 A:Title: cDNA cloning and characterization of human monocyte/macrophage serine esterase-1
 A:Reference number: A49816; MUID:91300111; PMID:2070086
 A:Accession: A49816
 A:Molecule type: mRNA
 A:Residues: 'G', 65-185, 'G', 187-361, 363-567 <ZSC>
 A:Cross-references: GB:M52973; NID:G36421; PID:CAA37147.1; PID:G1335304
 R:Riddler, P.W.; Richards, L.J.; Bowles, M.R.; Pond, S.M.
 Gene 108, 289-292, 1991
 A:Title: Cloning and analysis of a cDNA encoding a human liver carboxylesterase.
 A:Reference number: P80280; MUID:92084150; PMID:1748313
 A:Accession: P80280
 A:Molecule type: mRNA
 A:Residues: 114, 'H', 116-280, 'A', 282-300, 'IGNSYLTMYTRETQREBT', 318-336, 'R', 338-382, 'GSPT', 3
 A:Cross-references: GB:M65261; NID:G187028; PID:AAA83932.1; PID:G187029
 A:Experimental source: liver
 A:Note: differences between this sequence and other reports appear to be due to frameshif
 R:Kroetz, D.L.; McBride, O.W.; Gonzalez, F.J.
 Biochembiy 32, 11606-11617, 1993
 A:Title: Glycosylation-dependent activity of baculovirus-expressed human liver carboxyle
 A:Reference number: A48809; MUID:94032283; PMID:8218228
 A:Accession: I61085
 A:Structure: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-55, 'G', 57-361, 363-535, 'G', 537-567 <KRO1>
 A:Cross-references: GB:L07765; NID:G180949; PID:AAA35711.1; PID:G180950
 A:Accession: A48809
 A:Structure: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-17, 'A', 18-55, 'G', 57-535, 'G', 537-567 <R02>
 A:Cross-references: GB:L07764; NID:G180947; PID:AA16036.1; PID:G180948
 C:Genetics:
 A:Gene: GDB:CEB1; HMSN
 A:Cross-references: GDB:128044; OMIM:114835
 A:Map position: 16q13-16q22.1
 C:Superfamily: cholinesterase; cholinesterase homology
 C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; liver
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-145/Domain: propeptide #status predicted <PRO>
 F:50-553/Domain: cholinesterase homology <CH>
 F:146-567/Product: carboxylesterase #status experimental <MAT>
 F:564-567/Region: endoplasmic reticulum retention signal #status acyclical
 F:221,468/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.:	8,97e-79	Length:	567
Score:	1140.00	Matches:	248
Percent Similarity:	58.39%	Conservative:	79
Best Local Similarity:	44.29%	Mismatches:	197
Query Match:	1	Indels:	36
		Gaps:	9

US-10-023-515-1 (1-2158) x A41010 (1-567)

```

QY      129  TGGTCTTTTTCCTGATTCGACCCCTGTGGAGACAGACAGTGGGAAACTGGG  188
DB      2   TTPLeuArgAlaPheIleLeuAlaThrLeuSerAlaSerAlaAlaTrpGlnHis-----  19
QY      189  CCTTGTCTGAAGGCGCAGAGAGACACAGGCTGGATGATTCAGGCGCAAGCATC  248
DB      20  ProSeSerPro---ProValValaSerThrValHisGlyValaLeuGlyIleuValPheVal  38
QY      249  ACTGTGTGGAGACCCCTGTGCTGTGAACGTGTTCCTGAGATCCCTTGTGCTGCTCC  308
DB      39  SerLeuGlnGlyPheAlaGlnProValAlaIlePheLeuGlyIleProPheAlaValSerPro  58
QY      309  CGGCTGGAGATCCCTGGATTTAGAACCCGACGCTGTGATCCCTGGATTAATGGCA  368
DB      59  ProLeuGlyProLeuArgPheThrProGlnProAlaGluProTrpSerPheValIys  78
QY      369  GAAGCCACCTCTACCCCTAATTGTGCTCCAGAACTGAG-----TGGCTGCTCTTA  422
DB      79  AsnAlaThrSerTyrProMetCysThrGlnaSerProIyValaGlyIleuLeuSer  98
QY      423  GATCAACACATCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAACTGCTTAC  482
DB      99  GluLeuPheThrAsnArgIyValaIlePheLeuIleProLeuIyLeuSerGluAspCysLeuTyr  118
QY      483  CTGAACATCTATGCGCTGCGCCAGCGCATACAGGCTCCAGCTCCCGCTTGTGCTG  542
DB      119  LeuAsnIleTyrThrProAlaPheLeuThrIyValaIlePheLeuIleProValMetValTrp  138
QY      543  TTCCCAAGAGGTGCTTCAAGACTGGCTCAAGCTTCATCTTGTGATGGTCCGCTGAGT  602
DB      139  IleHisGlyGlyIyLeuMetValGlyAlaIleSerThrTyrAspGlyLeuAlaIleuAla  158
QY      603  GCGTAATGAGAGCGTGTGTGTGTGCTCCAGTACCGGCTAGAAATTTGGTTTCTTC  662
DB      159  AlaHisGluAsnValValaValaThrIleGlnTyrArgLeuGlyIleTrpGlyPhePhe  178
QY      663  ACACATGGAGTACAGACTGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGCTG  722
DB      179  SerThrGlyArgIleHisSerArgIyAsnTrpGlyIleuAspGlnValaIleuAlaIleu  198
QY      723  TCCCTGGTTCGAAGAAATCAAGTTCTTCGCTGGAGACCCCAAGCTGTGACCATTTT  782
DB      199  ArgTrpValGlnaPheAsnIleAlaSerPheGlyIyAsnProGlySerValThrIlePhe  218
QY      783  GCGCAGTCCGCGGAGCATAGTGTTCATAGTCTTATAGTGTCTCCCAAGCGCAAGGC  842
DB      219  GlyIleuSerAlaGlyIyGlyIleuSerValSerValIleuSerProLeuAlaIyAsn  238

```

```

QY      843  TTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTACCTGAGGCCCAT  902
DB      239  LeuPheHisArgAlaAlaSerGlySerGlyValaIleuThrSerValLeuValIyValys  258
QY      903  GATTTATGAGAG-----AGTGAAGACCTGCAAGGTGGTTCACATTTCTGTGTAAACAT  956
DB      259  GlyaPheValIyPheProLeuAlaGlnIleIleAlaIleThrIle-----GlyCysValys  275
QY      957  GCGTCAAGACTGAGAGCCCTGAGTGGCTGAGTCAAGAACAAACCTCCAGAGAGTGTG  1016
DB      276  ThrThrThrSerAlaValaMetValHisCysLeuArgGlnIyIyThrGlnGlnGlnIleu  295
QY      1017  ACCCTCAGCCAGAAACAAACCTTTACT-----GlyCysValys-----  1046
DB      296  GluThrThrIleuIyMetIyPheLeuSerIleuAspLeuGlnGlyAspProArgGlySer  315
QY      1047  -----CGAGTGTGATGTGAGTCTTTCTTTCTTATGAGCCCTTGAATCTA  1091
DB      316  GlnProLeuLeuGlyIyThrValIleAspIyMetLeuLeuIyIyThrProGlnGlnIleu  335
QY      1092  TTGTCTCAAGAAAGCATTTAAAGCAATTCCTTCATCAATCGAGTCAATTAACAGAGT  1151
DB      336  GlnIleGlnaArgaPheHisIleThrValProIyMetValGlyIleAsnIyGlnGlnIle  355
QY      1152  GAGTCTCTGCTGCTTATGAGAGAGCTCTGAGATCTCTAGTGGCTCC-----AACAG  1205
DB      356  GlyTrpLeuIleProMetGlnIleuMetSerTyrProLeuSerGlnGlyIleuAspGln  375
QY      1206  TCCCTTGCCTTCATCTGATTAACAAC-----ATCCGACATCCCGCTCAAGTAT  1256
DB      376  LysThrAlaMetSerIleuLeuTrpIySerTyrProLeuValIyIleAlaIyGlnIleu  395
QY      1257  TTGACCTGTGGGTAAATGAAATCTTCATGACAGACACTCCCTGACTGAATATCGAGAC  1316
DB      396  IleProGlnAlaThrGlnIyIyTyrLeuGlyIyThrAspAspThrValIyIyValysAsp  415
QY      1317  AGTCTTGTGAGATGTGCTGTGATGATGTCTTGTGCTGCTGCTGCTGACATGACATGCA  1376
DB      416  LeuPheLeuAspLeuIleAlaAspValMetPheGlyValaProSerValIleValaIleArg  435
QY      1377  TATCAACAGAAATGCTGTGTCACCTGTCTTACTTATGAGTTTGAGACCGGCTCAAGTC  1436
DB      436  AsnHisIleArgAspAlaGlyAlaProThrIyMetTyrGluPheGlnIyArgProSerPhe  455
QY      1437  TTGGAAGACAGAAAGCGGCTTTGTCAAGCGCAGACGATGAATCCGCTTGTG  1496
DB      456  SerSerAspMetCysProIyIyThrValIleGlyAspHisGlyAspIyIleuPheSerVal  475
QY      1497  TTGGTGTGTGCTTCTTGAAGGGGACATTTGTTATGTTGGAAGAGCCAGAGAGAGAG  1556
DB      476  PheGlyAlaProPheLeuIyS-----GlnGlyAlaSerGlnGlnIleu  489
QY      1557  AAGTTACTAGACCGGAAAGATGATGAATACTGGGCTAATTTGCTGGAACCGGAAATCT  1616
DB      490  IleArgLeuSerIyMetValMetIyIyPheThrAlaAsnPheAlaArgAsnIyAsnPro  509
QY      1617  AATGGAAAGACCTGTCTGTGTGCGCAGCTTATATCTGATGAGACAGATCCAGAGCT  1676
DB      510  AsnGlyGlyGlyIleuProHisIyThrProGlnIyIyIyAsnGlnIyIyIleuGlnIle  529
QY      1677  GACTTGAACATGAGCCTCGACAGAGACTCAAAAGAACCGGCGGTAATTTTGAACAGC  1736
DB      530  GlyAlaAsnThrGlnAlaIleGlnIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy  549

```

RESULT 10

S62788
 carboxylesterase (EC 3.1.1.1) ES-4 precursor, liver - rat
 N:Alternate names: hydrolase B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Oct-1996 #sequence revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S62788; S51203; S49257
 R:Robbt, M.; van Schaftingen, E.; Beaufay, H.
 Biochem. J. 313, 821-826, 1996

A:Title: Cloning and sequencing of rat liver carboxylesterase ES-4 (microsomal palmitoyl
A:Reference number: S62788; MUID:96190723; PMID:8611161
A:Accession: S62788
A:Molecule type: mRNA
A:Residues: 1-561 <ROB>
A:Cross-references: UNIPROT:064573; EMBL:X81825; NID:G550417; PIDN:CAA57419.1; PID:G5504
A:Experimental source: liver
R:Morgan, B.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A:Title: Purification and characterization of two rat liver microsomal carboxylesterases
A:Reference number: S51202; MUID:95077430; PMID:7986098
A:Accession: S51203
A:Molecule type: protein
A:Residues: 19-48 <MOR>
A:Experimental source: liver
R:Yan, B.; Yang, D.; Brady, M.; Parkinson, A.
J. Biol. Chem. 269, 29688-29696, 1994
A:Title: Rat kidney carboxylesterase. Cloning, sequencing, cellular localization, and re
A:Reference number: A55304; MUID:95050819; PMID:7961958
A:Accession: A55304
A:Molecule type: mRNA
A:Residues: 1-6, 'P', 8-59, 'P', 61-212, 'A', 214-252, 'T', 254-309, 'T', 312-341, 'N', 343-424, 'P'
A:Cross-references: GB:U10697; NID:9562007
A:Note: the sequence in Genbank entry RNU10697, release 107, (PID:G562008) has the codon
R:Robbi, M.; Beaufay, H.
submitted to the EMBL Data Library, September 1994
A:Reference number: S49257
A:Accession: S49257
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <RO2>
A:Cross-references: EMBL:X81825; NID:G550417; PIDN:CAA57419.1; PID:G550418
A:Function: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
A:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
P:1-18/Domain: signal sequence #status predicted <SIG>
P:19-561/Product: carboxylesterase ES-4 #status experimental <MUT>
P:50-551/Domain: cholinesterase homology <CHS>
P:221,466/Active site: Ser, His #status predicted
P:301/Binding site: carbohydrate (Asn) (covalent) #status predicted
Alignment Scores:
Pred. No.: 1,07e-78 Length: 561
Score: 1139.00 Matches: 243
Percent Similarity: 58.04% Conservative: 82
Best Local Similarity: 43.39% Mismatches: 195
Query Match: 29.42% Indels: 40
DB: 2 Gaps: 10
US-10-023-515-1 (1-2158) x S62788 (1-561)
QY 138 TTCCTGATTTCTCAGCCCTGTGGGACACAGACAGTGGGAAACCTGGCCCTTGTCT 197
DB 5 PheLeuIleLeuValSerLeuAlaThrCysValValTyrGly-----AsnProSerSer 22
QY 198 GAAAGGCCACAGAGAACACCGAGCGCTGGATTCAGGGCAAGACATCACTGTGCTG 257
DB 23 Pro---ProValValAspThrThrLysGlyLysValLeuGlyLysTyrValSerLeuGlu 41
QY 258 GGAAGCCCGTGCCTGTAAGAGTGTTCCTGGAGTCCCTTGTGCTGCTCCCGCGTGGGA 317
DB 42 GlyValThrGlnSerValAlaValPheLeuGlyValPhePheAlaLysProProLeuGly 61
QY 318 TCCCTGCGATTTACGAACCCGAGCCTGCATGCCTGGAGATTAATTCGAGAAAGCCACC 377
DB 62 SerLeuArgPheAlaLysProGlnProAlaGluProTyrSerPheValLysAsnThrThr 81
QY 378 TCCCTACCTTAATTTGTGCTCCAGAACTCAAGATGG-----CTGCTCTTA 422
DB 82 ThrTyrProProMetCysSerGlnAspAlaAlaLysGlyGlnArgMetAsnAspLeuLeu 101
QY 423 GATCAACACATGCTCAAGTGCATTAACCGAAATTCGAGTGTCAAGAACCTGCCTAC 482

DB 102 ThrAsnArgLysGlyLysLeuIleHis-----LeuGluPheSerGluAspCysLeuTyr 118
QY 483 CTGAACATTTATGCGCTGCCCAACCCGATACAGGCTCCCAAGCTCCCGCTTGTGTGG 542
DB 119 LeuAlaThrIleTyrThrProAlaAspPheThrLysAsnSerArgLeuProValIleTyr 138
QY 543 TTCGAGAGAGTGCCTTCAAGCTGAGCTCAGCTCCATCTTTGATGGTGCCTGCGCTG 602
DB 139 IleHisGlyValGlyMetThrLeuGlyGlyAlaSerThrTyrAspGlyArgValLeuSer 158
QY 603 GCGTATGAGACGTGCTGTTGTGCTGCTCAATCCGCTAGAAATATTTGGTTCTTC 662
DB 159 AlaTyrGluAsnValValValValAlaIleGlnTyrArgLeuGlyLysTyrGlyPhePhe 178
QY 663 ACCACATGGGATCAGATCTCCGCGGAACTGGGCTTCAAGACCAAGCTGCTGTCTG 722
DB 179 SerThrGlyAspGluHisIleSerArgLysAsnTyrGlyHisLysLeuAspGluValAlaIleLeu 198
QY 723 TCCGCGTCCAGAAACATCGAGTTCTTCGGTGGGAGCCCGAGCTGTGACCATCTTT 782
DB 199 HisThrValGlnAspAsnIleAlaAsnPheGlyGlyAspProGlySerValThrIlePhe 218
QY 783 GCGAGTCCGCGGAGCCATTAAGTGTTCCTTAATCTGTCTCCATGCGCAAGGC 842
DB 219 GlyGlnSerAlaGlyGlyPheSerValSerValLeuValLeuSerProLeuThrLysAsn 238
QY 843 TTATTCACAAAGCCATCAGAGAGTGGGGTGCATATCCCTTACCTGAGGCCCAT 902
DB 239 LeuPheHisArgAlaIleSerGlnSerGlyValAlaPheLeuProGlyLeuLeuThrLys 258
QY 903 GATTATGAGAGAGAGAGACCTGCAGCGTGTGACATTTCTGTGGTAAACAATGGCTCA 962
DB 259 AspValAlaGlyProAla-----AlaLysGlnIleAlaAspMetAlaGlySerGluThrThr 276
QY 963 GACTGTAGGCCCTCTGAGTGCCTTAGAGCAAAACCTCCAGAGAGCTGTGACCCCT 1022
DB 277 ThrSerAlaIleIleValHisCysLeuArgGlnLysThrGluGluGluLeuGluIle 296
QY 1023 AGCCAGAAACAAG----- 1037
DB 297 MetLysLysMetAsnLeuIleLysLeuSerSerGlnArgAspAsnLysGlySerTyrHis 316
QY 1038 TCTTTCACGTCGAGTGTGATGATGCTTTCTTCCATAGAGCTCTAGATCATGTGCT 1097
DB 317 PheLeuSerThrValValAspAsnValValLeuProLysAspProLysGluLeuAla 336
QY 1098 CAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATACCAAGAGTGTGCTTC 1157
DB 337 GlyLysAsnPheAsnThrValProTyrIleValGlyIleAsnLysGlnGluCysGlyTyr 356
QY 1158 CTGCTGCT-----ATGAGAGAGCTCCTGAGATCTTCAGTGGCTCCAAAGTCCCTT 1211
DB 357 LeuLeuProThrMetMetC1yPheValProAlaAspValGlu---LeuAspLysLysMet 375
QY 1212 GCCCTTCATCTGATCA-----AACATCTGCACATCCCGCTCAGATTTGAC 1262
DB 376 AlaIleThrLeuLeuGlnLysPheAlaSerLeuTyrGlyLysProGluAspIleIlePro 395
QY 1263 CTGTGCGCTAATGATCACTTCATGACAAAGCACTCCCTGACTCAATTCGAGAGCTTT 1322
DB 396 ValAlaIleGlyLysTyrArgLysSerAspAspSerIleLysIleArgAspGlyIle 415
QY 1323 CTGACCTGCTTGAAGATGTGTTTGTGCTCCCTGACATGATCAACAGCTGATATC 1382
DB 416 LeuAlaPheIleGlyAspValSerPheSerIleProSerValMetValSerArgAspHis 435
QY 1383 AGAGATCTGTGACACCTGTCACTTCTATGAGTTTGCGACCGGCTCAGTGGCTTGA 1442
DB 436 ArgAspAlaGlyAlaProThrTyrMetCyrGlnTyrGlnTyrTyrProSerPheSerSer 455
QY 1443 GACAGAAAGCCGCTTTGTCAAAACCCGACCAAGCTGATGAATCGCTTGTGTGCT 1502

Db 456 ProGlnArgProLeuHisValValGlyAspHisAlaAspLeuTyrSerValPheGly 475
Qy 1503 GGTGGCTTCCTGAAGGGGACATTTGTATGTTCCAGAGGCCAGAGAGAGAGATTGA 1562
Db 476 AlaProIleLeuArg-----AerGlyAlaSerGluGluGluIleLeu 489
Qy 1563 CTGAGCCGAGAGATGATGAAATACTAGGCTACCTTGGTCCGAAOCGGGAGATCCCTAAAGG 1622
Db 490 LeuSerLeuMetValMetLeuPheThrPAlaAsnPheAlaArgAsnGlyAsnProAsnGly 509
Qy 1623 AACGACCTGTCTCTGTGGCCAGCTTAAATCTGACTGACGACTCTTCAGCTGACTTGG 1682
Db 510 ArgGlyLeuProHisAlaTrpProGlnIleTyrAspGlnGluGluTyrLeuGlnIleGlyAla 529
Qy 1683 AACATGAGCTTCGAGACAGACTCAAGAGACCGGGGGGTGATTTTGGACACAGACATC 1742
Db 530 ThrThrGlnGlnSerGlnArgLeuLeuValGluGluValAlaPheThrGlnLeuLeu 549

RESULT 11
A55281
carboxylesterase (EC 3.1.1.1) egaayn - mouse
N/Alternate names: beta-glucuronidase endoplasmic reticulum-targeting protein; esterase-
C/Species: Mus musculus (house mouse)
C/Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C/Accession: A55281
R/Ovnc: M.; Swanik, R.T.; Fletcher, C.; Zhen, L.; Novak, E.K.; Baumann, H.; Heintz, N.;
Genomics 11, 956-967, 1991
A/Title: Characterization and functional expression of a cDNA encoding egaayn (esterase-
A/Reference number: A55281, MUID:92147141, PMID:1783403
A/Accession: A55281
A/Status: preliminary
A/Molecule type: mRNA; protein
A/Residues: 1-562 <OVN>
A/Cross-references: UNIPROT:Q64176; GB:SB0191; NID:9244727; PIDN:AB21335.1; PID:9244728
A/Note: sequence extracted from NCBI backbone (NCBIN:80191, NCBITP:80194)
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; endoplasmic reticulum
F/51-552/Domain: cholinesterase homology <CH>
F/222.467/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,17e-78 Length: 562
Score: 1138.50 Matches: 240
Percent Similarity: 58.52% Conservative: 76
Best Local Similarity: 44.44% Mismatches: 195
Query Match: 29.40% Indels: 29
Gaps: 6

US-10-023-515-1 (1-2158) x A55281 (1-562)

Qy 186 GGGCCCTTCGTGAAGGGCCACAGAGAGAACACAGGCTGGATTCAGAGGACAGCA 245
Db 19 GlnHisProSerSerProProMetValAlaThrValGlnGlyValLeuGlyTyrSer 38
Qy 246 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCTGGAGCTCCCTTGGCT 305
Db 39 IleSerLeuGluGlyPheThrGlnProValAlaValPheLeuGlyValProPheAlaLeu 58
Qy 306 CCCCCCTGGGATCCCTGCGATTACGAACCCGACGCTGCATGCGCCCTGGATTAACCTG 365
Db 59 ProProLeuGlySerLeuArgPheAlaProProGlnProAlaGluProTropSerSerVal 78
Qy 366 CGAGAGCCACCTCTCAATTTGTCCTCCAGAACCTCAAGACAGAGCTGCTCTTGAT 425
Db 79 LysAsnAlaThrSerTyrProProMetCysPheGlnAspProValThrGlyGlnIleVal 98
Qy 426 CAACACATGCTCAAGGTGATTAACCGAATTCGAGTG-----TCAGAGACTGCTC 479
Db 99 AsnAspLeuLeuThrAsnArgGlyGluValIleProLeuGlnPheSerGluAspCysLeu 118
Qy 480 TACCGAAGATCTATGCGCTTCGCCACGCGCATACAGGCTCCAGCTCCCGTCTTGGT 539
Db 119 TyrLeuAsnIleTyrThrProAlaAspLeuThrLysSerAspArgLeuProValMetVal 138

Qy 540 TGGTCCAGAGAGTGCCTTCAAGACTGGCTCAGGCTCCATCTTGATGGTCCGCTG 599
Db 139 TrpIleHisGlyGlyGlyLeuValLeuGlyGlyAlaSerThrTyrAspGlyLeuValIleu 158
Qy 600 GCTGCTATGAGAGAGCTGTGGTGTGTGCTGTCTGCTCAATCCGCGCTAGAAATATTTGGTTG 659
Db 159 SerThrHisGluAsnValValValValIleGlnTyrArgLeuGlyIleTrpGlyPhe 178
Qy 660 TTCACACATGGGATCAGAGTCTCCGGGAACCTGGGCTTCAGAGACAGAGTGGCTGT 719
Db 179 PheSerThrGlyAspGluHisSerArgGlyAsnTrpGlyHisLeuAspGlnValAlaAla 198
Qy 720 CTGTCTGGGTCCAGAGAACATCGATGTTCTTGGTGGAGCCGACGCTGTGACATC 779
Db 199 LeuHisTrpValGlnAspAsnIleAlaLysPheGlyGlyAspProGlySerValThrIle 218
Qy 780 TTTGGCGAGTCCGCGGAGCCATTAAGTGTTCATGCTTAATATGTCTCCATGGCCAA 839
Db 219 PheGlyGluSerAlaGlyGlyGlySerValSerValLeuValLeuSerProLeuAlaLys 238
Qy 840 GGTATTATCCAAAGACCATGATGAGAGTGGGGGCGCATCATCCCTTACAGAGCTGTGACC 899
Db 239 AsnLeuPheGlnArgAlaIleSerGluSerGlyValAlaLeuThrAlaGlyLeuValLys 258
Qy 900 CATGATTATGAGAGAGTGAAGACCTCGAGGTGTCACATTTCTGTGGTAACAATGCG 959
Db 259 LysAsnThrArgProLeuAlaGluValIleAlaValIleSerGlyCysLysAsn----- 276
Qy 960 TCAGACTGTGAGGCGCTGATGAGTGGCTGATGAGGACAAACCCCTCAAGAGCTGTGACC 1019
Db 277 ThrThrSerAlaAlaMetValHisCysLeuArgGlnLysThrGluGluGluLeuGly 296
Qy 1020 CTCAGCCAGAAACAAAGCTCTTC----- 1043
Db 297 ThrThrLeuLysLeuAsnLeuPheLysLeuAspLeuHisGlyLysSerArgGlnSerHis 316
Qy 1044 -----ACTCGAGTGTGATGATGCTGTTCTTCTTCAATGACGCTTCAATGTTG 1094
Db 317 ProPheValProThrValIleLeuAspGlyValLeuLeuProLysMetProGluGluIleLeu 336
Qy 1095 TCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCATGAGTCAATTAACAGAGTGTGGC 1154
Db 337 AlaGluLysAsnAsnAsnThrValProTyrIleValGlyIleAsnLysGlnGluPheGly 356
Qy 1155 TTCTGTGCTGCT---ATGAAGAGGCTCTTGATGATCTCGAGTGCCTCAAGAGCTCCCTT 1211
Db 357 TrpIleLeuProThrMetMetAsnTyrProProSerAspValLysLeuAspGlnMetThr 376
Qy 1212 GCCCTCATGTGATACAAAC-----ATCCGACATATCCGCTCAGATATTGCAC 1262
Db 377 AlaMetSerLeuLeuLysSerSerPheLeuLeuAsnLeuProGluAspAlaIleAla 396
Qy 1263 CTGTGAGCTTAAGATACTTCATGATGACAGACATCCCTGATGAAATCCGAGACAGTCTT 1322
Db 397 ValAlaIleGluTyrLeuArgAspLysAspTyrThrGlyArgAsnLysAspGlnLeu 416
Qy 1323 CTGAGCTGTGAGATGTTGTTTGTGTGGTCCCTGCACTGATACAGCTGATATAC 1382
Db 417 LeuGlnLeuIleGlyAspValAlaPheGlyValProSerValIleValSerArgIyHis 436
Qy 1383 AAGATGCTGTGACACTGTCTAATCTTATGATGTTGGACAGCGGCTCAAGTCTTGAA 1442
Db 437 ArgAspAlaGlyAlaProThrTyrMetTyrGluPheGlnTyrSerProSerPheSerSer 456
Qy 1443 GACAGAGAGCGGCTTTGTCAAGCCGACAGCTGATGAAGTCCGCTTGTGTTCGT 1502
Db 457 GluMetLysProAspThrValValGlyAspHisGlyAspGluIleTyrSerValPheGly 476
Qy 1503 GGTGCTTCCTGAAGGGGAGCATTTGTATGTTCAAGAGGACGAGAGAGAGAAATTGA 1562
Db 477 AlaProIleLeuArgGly-----GlyThrSerGluGluGluIleAsn 490

```

QY 1563 CTGAGCCGGAAGATGATGAAATTAAGTGGGCTTGTGCTGCAACCGGGAATCTTAATGGG 1622
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 491 LeuSerIyWMeMetMeMetLyBheTrrAlaAsnPhelAlaArgAsnGlyAsnProAsnGly 510
QY 1623 AACGACCTGTCTGTGGCCAGCTTAATTAATGACTGAGACGATCTCCAGCTGCACTTG 1682
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 511 GlnGlyLeuProHlaTrrProGluIuTyArgGlnGlyGlnGlyLeuGlnIleGlyAla 530
QY 1683 AACATGACCTCGGACAGACGATCAAGACCGCGGGTGGATTTTGGACGACGACATC 1742
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 531 ThrThrGlnAlaGlnIlyLeuIlyGlnIlyGlnIlyValAlaPheTrrThrGlnIleLeu 550

RESULT 12
A31584
carboxylesterase (EC 3.1.1.1) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-May-1990 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004
C:Accession: A31584
R:Long, R.M.; Satoh, H.; Martin, B.M.; Kimura, S.; Gonzalez, F.J.; Pohl, L.R.
Biochem. Biophys. Res. Commun. 156, 866-873, 1988
A:Title: Rat liver carboxylesterase: cDNA cloning, sequencing, and evidence for a multig
A:Reference number: A31584; MIMD:89050119; PMID:2973315
A:Accession: A31584
A:Molecule type: mRNA
A:Residues: 1-540 <LON>
A:Cross-references: UNIPROT:P10959; GB:M20629; GB:X13587; NID:G203279; PIDN:AAA40871.1;
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein
F:1-9/Domain: signal sequence #status predicted <Sig>
F:10-540/Product: carboxylesterase #status predicted <Mnt>
F:41-529/Domain: cholinesterase homology <CHS>
F:70-265,266,293,366,467/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:212,444/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1.79e-78 Length: 540
Score: 1136.00 Matches: 237
Percent Similarity: 59.96% Conservative: 91
Best Local Similarity: 43.33% Mismatches: 185
Query Match: 29.34% Indels: 34
DB: 2 Gaps: 9

US-10-023-515-1 (1-2158) x A31584 (1-540)
QY 141 CTGATTCCTCCAGCCCTGTGGGACACAGACGTTGGGAAAACTGGCCTTCTGCTGA 200
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 2 LeuAlaValCySProIleTrrGlyHis-----ProSerSerPro 14
QY 201 GGGCCACAGAGAACACAGCGCTGGAGTGAATTCAGGGACAGCACTGCTGCTGGGA 260
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 15 ---ProValValAspThrThrIyGlyIyValIleuGlyIyValIySerIleuGlnGly 33
QY 261 AGCCCTGTGCTGTGAAGCGTGTCTCGAGAGTCCCTTGTGCTGCTCCCGCGTGGATCC 320
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 34 PheThrGlnProValAlaValPheIleuGlyValPhePheAlaIyProProIleuIySer 53
QY 321 CTGGGATTTACGAAACCGCAGCTGCATCGCCTGGGATTAATTCGAGAAACCACTTCC 380
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 54 LeuAlaPheAlaIyProGlnIyProIleuGlnIyProIleuSerPheValIyAsnThrThrThr 73
QY 381 TACCTAATTTGTGCTCCAGAAC-----TCAGAGTGGTGTGCTC 419
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 74 TyrProProMetCySPerGlnAspIyValIyValIyIyValIyLeuIleuAlaAspMetLeuSer 93
QY 420 TTAGATCAACATGCTCAAGTGTGATTATCCGAAATTCGAGAGTGTGAGAGATGCTCTC 479
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 94 ThrGlyIyGlnSerIlePheIleuGlnIyPhe-----SerGlnAspCySerIleu 108
QY 480 TACCTGAACATCTATGCGCTGCCACGCGCATATACGCTCCAGACTCCCGCTTGGTG 539
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 109 TyrIleuAlaIyIySerProIleuAspIyLeuThrIyIyAsnSerArgIleuProValIleuVal 128
QY 540 TGGTCCAGAGAGGTCCTTCAAGCTGGCTCAGGCTCATTCTTGGAGGGTCCGCGCTG 599

```

```

Db 129 TrpIleHlaGlyGlyLeuIleIleIleGlyAlaSerProIySerGlyLeuAlaIleu 148
QY 600 GCTGCGCTATGAGACGTCGTGTGTGTCGTCGACGTAACCGGCTAGAAATATTGTTGCTTC 659
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 149 SerAlaHlaGlnAsnValValValThrIleGlnIyIyArgIyGlnGlyPheGlyIyIleu 168
QY 660 TTACACCATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGGTGGTGGCT 719
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 169 PheSerThrGlyAspGlnHisSerArgIyAsnThrAlaHisIleuAspGlnIleuAlaIa 188
QY 720 CTGTCTCGGCTCCAGAGAACATCAGTCTTGGTGGGACCCGACCTCTGTGACATC 779
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 189 LeuArgTrrValGlnAspAsnIleAlaAsnPheGlyIyAsnProAspSerValThrIle 208
QY 780 TTGGCGGAGTCCCGGGAGCCATAGTGTCTTAATCTTAATCTTCCCATGGCCAA 839
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 209 PheGlyGlnSerHlaGlyIyValSerValSerAlaIleuValIleuSerProIleuAlaIyS 228
QY 840 GGCTATTCCCAAAAGCCATCATGAGAGTGGGGGCGCATCATCTTACCTGAGAGCC 899
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 229 AsnIleuPheHlaArgAlaIleSerGlnSerGlyValIleuThrThrAsnIleuAspIyS 248
QY 900 CATGATTATGAGAGAGTGAAGACCTGCAGGTGTGACATTTCTGTGTAAACAATGCG 959
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 249 IyAsnThrGlnAlaVal-----AlaGlnMetIleAlaThrIleuSerGlyCyAsnAsn 266
QY 960 TCAGACTGTGAGGCGCTCTGAGTGTGAGAGACAAACCTCCAAAGAGTGTGACAC 1019
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 267 ThrSerSerAlaAlaIleMetValGlnCyIleuArgIyIyThrGlnAlaGlnIleuGln 286
QY 1020 CTCAGCCCAAAA-----ACAAAGCTTTCATCGAGGTGTGATGATGCTTCTTCTCT 1073
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 287 LeuThrValIyIyLeuAspAsnThrSerMetSerThrValIleAspGlyValIyLeuPro 306
QY 1074 AATGAGCTCTAGATCTAATTTGCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGA 1133
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 307 IySerThrProGlnIyGlnIleuThrGlnIySerPheAsnThrValProTyIleValGly 326
QY 1134 GTCAATTAACCAAGAGTGGCTCTGCTGCTGCTATGAAGAGGCTCCAGATCTCAAT 1193
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 327 PheAsnIyGlnGlnIyPheGlyIyTrpIleIleProThrMetMetGlyAsnIleuSerGln 346
QY 1194 GGC---TCCAAAGTCCCTGCGCTCCATCATGATCAAAACATC-----CTGCAC 1241
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 347 GlyArgMetAsnGlnIyMetAlaIleSerSerPheIleuIyAspGlnSerProAsnIleuAsn 366
QY 1242 ATCCGCGCTCAGTATTTGACCTTGCTGCTAATGATCTTCATGACAGACATCCCTG 1301
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 367 IleSerGlnSerValIleProAlaIleIleGlnIySerIyLeuArgGlyThrAspAspPro 386
QY 1302 ACTGAATCCAGACAGATCTTCTGAGCTTGTGAGATGTGTGTTTGTGTGCTCCCTGA 1361
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 387 AlaIyIyIyGlnIleuIleuAspMetPheSerAspValPhePheGlyIyIleProAla 406
QY 1362 CTGATCAAGCTCATATCAAGAGATGCTGTGACCTGTCTAATTCATAGAGTTGG 1421
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 407 ValIleuMetSerArgSerIleuArgAspAlaGlyAlaProThrIyMetIyIleuGln 426
QY 1422 CACCGGCTCAGTGTGTTGAAGACAGAGCGGCTTTGTCAAGCGGACCAAGCCTGAT 1481
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 427 TyrArgProSerPheValIleSerAspGlnArgProGlnThrValGlnIyAspHisIleIyAsp 446
QY 1482 GAAGTCCGCTTGTGTGTGCTGTGCTCTTCTGAAGGGGACATTTGATGTTGCAAGGA 1541
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 447 GlnIlePheSerValPheGlyIyThrProPheIleuIyS-----GlnIyIy 460
QY 1542 GCCACGAGAGAGAGAAATTAATGACCGGAAAGATGAATATCTGGGCTACCTTGTCT 1601
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 461 AlaSerGlnGlnIyGlnThrAsnIleuSerIyIleuValIleuIyPheTrrAlaAsnPheAla 480
QY 1602 CGAACCGGGAATCTTAATGGAACGACCTGTCTGTGCGGACGCTTAATATCTGACTGAG 1661
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

Db	481	AtgaannglyAsnProAsnnglyGluLeuProHisTrpProLysTrpAspGlnyGlu	500
QY	1662	CAGTACCTCCAGCTGACTGGAACATGAGCCTCGACAGAGACTAAAGAACCGGGGTG	1721
Db	501	GlyTrpLeuGlnIleGlyAlaThrThrGlnGlnAlaGlnLysLeuSerGlyGluVal	520
QY	1722	GATTTTGGACCAGCACCATC	1742
Db	521	AlaPheTrpThrGluLeuLeu	527

QY 1611 AATCTAATGGGAACGACCTGTCTCTGTCGACGTTATATCTGACTGACGATACCTC 1670
|||||
Db 506 AaenProAenAlaArgGlyLeuProHisTrpProGlnTrpAspGlnLysGlnLysTrpLeu 525
|||||
QY 1671 CAGCTGACCTTGACATGAGCTCTCGACAGACGCTCAAGAAACCGCGGTGATTTTGG 1730
|||||
Db 526 GlnIleGlyAlaThrThrGlnInSerGlnArgLeuLysAlaGlnLysAlaIlePheTrp 545
|||||
QY 1731 ACCGACCACTC 1742
|||
Db 546 ThrGlnLeuLeu 549
|||
RESULT 14
JC2447
carboxylesterase (EC 3.1.1.1) ES-3 precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 29-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C/Accession: JC2447; S23462
R/Robbi, M.; Beaufay, H.
Biochem. Biophys. Res. Commun. 203, 1404-1411, 1994
A/Title: Cloning and sequencing of rat liver carboxylesterase ES-3 (EGASYN).
A/Reference number: JC2447; MUID:95032008; PMID:7945287
A/Accession: JC2447
A/Molecule type: mRNA
A/Residues: 1-561 <ROB>
A/Cross-references: UNIPROT:063108; GB:X81395; NID:9550146; PIDN:CAA57158.1; PID:9550147
A/Experimental source: Liver
B/Ur. J. Biochem. 206, 801-806, 1992
B/Medda, S.; Proia, R.L.
A/Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
A/Reference number: S23460; MUID:92299008; PMID:1606962
A/Accession: S23462
A/Status: preliminary; translation not shown
A/Molecule type: mRNA
A/Residues: 503-554, 'R', 556-561 <MED>
A/Cross-references: EMBL:X65295; NID:957557; PIDN:CAA46390.1; PID:957558
C/Superfamily: cholinesterase; cholinesterase homology
C/Keywords: carboxylic ester hydrolase; glycoprotein
F/1-18/Domain: signal sequence #status predicted <SIG>
F/19-561/Product: carboxylesterase ES-3 #status predicted <MAT>
F/50-551/Domain: cholinesterase homology <CHB>
F/79,107,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/221,466/Active site: Ser, His #status predicted
Alignment Scores:
Pred. No.: 1.78e-77 Length: 561
Score: 1123.00 Matches: 245
Percent Similarity: 57.07% Conservative: 86
Best Local Similarity: 42.24% Mismatches: 193
Query Match: 29.00% Gaps: 56
Db: 2 Indels: 11
US-10-023-515-1 (1-2158) x JC2447 (1-561)
QY 132 TGCCTTTTCGATTCCTCCAGCCCTGTTGGACACACAGTGGGAAAACTGGGCT 191
|||
Db 2 CysLeuTrpAlaLeuIleLeuValPheLeuAlaAlaPheThrAlaGly-----GlnHis 19
|||
QY 192 TCTGTGAAGGCGCCAGACAGACAGAGGTGGATGATTCAGGGCAAGCACTCACT 251
|||
Db 20 ProSerSerLeuProValAlaPheThrLeuGlnGlyLysValLeuGlyLysTrpValSer 39
|||
QY 252 GTGCTGGAGACCTGTGCTGTGTAACGTGTTCTCGAGTCCCTTTGCTGCTCCCGC 311
|||
Db 40 LeuGlnGlyPheThrGlnProValAlaValPheLeuGlyValProPheAlaLysProPro 59
|||
QY 312 CTGGATCCCTCGGATTAACGAACCCGACGCTGATGCTCCCTGGGATTACTGGAGAA 371
|||
Db 60 LeuGlnSerLeuArgPheAlaProGlnProAlaGlnProTrpSerPheValLysAsn 79
|||
QY 372 GCCACCTCTCAATTTGTCCTCCAGAAC----- 404
|||

Db 80 ThrTrpSerTrpProPheMetCysSerGlnAspProValAlaGlyGlnIleValAsnAsp 99
QY 405 -----TCAGAGTGGCTGCTCTTATGATCAACAGATGCTCAAGGTGATTAACCGAAATTC 458
|||
Db 100 LeuLeuTrpAsnTrp-----GlnGlnAsnIleSerLeuGlnPhe----- 112
|||
QY 459 GGAGTGTCAAGAGACTGCTTACCTGAACATCTATGCGCTCCCGACGCGATACAGGC 518
|||
Db 113 -----SerGlnAspCysLeuTrpLeuAsnIleTrpThrProAlaAspLeuThrLysArg 130
|||
QY 519 TCAGAGCTCCCGCTTGTGTGTGTTCCAGAGAGTGCCTTAAAGACTGGCTCAGCTTC 578
|||
Db 131 AspArgLeuProValMetValTrpIleHisGlyGlyGlyLeuValLeuGlyValAsp 150
|||
QY 579 ATCTTGAATGGGTGCGCCCTGCTGCTGCTGATGAGAGAGTGGTGTGGTGTGCTCAAGTAC 638
|||
Db 151 ThrTrpAspGlyLeuAlaLeuSerThrHisGlnAsnValValValValIleGlnTrp 170
|||
QY 639 CGGCTAGGAATATTTGGTTTCTTACACACATGAGGATGACATGCTCCGCGAAGTGGGCC 698
|||
Db 171 ArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGlnHisSerArgLysTrpGly 190
|||
QY 699 TTCAAGACCAAGTGGCTGCTGTCTGCTGCTGCTCAAGAAACATGATGATTTCCGTGG 758
|||
Db 191 HisLeuAspGlnValAlaAlaLeuHisTrpValGlnAspAsnIleAspAsnIleGly 210
|||
QY 759 CACCCCACTCTGTGACATCTTTGGCGAGTCCGCGGAGCCATAGGTTTCTAGTCTT 818
|||
Db 211 AspProGlnSerValThrIlePheGlyGlnSerAlaGlyGlyGlnSerValSerValLeu 230
|||
QY 819 ATATGCTCTCCAGTCCCAAGAGCTTATTCACAAAGCATGATGAGAGTGGGTGGCC 878
|||
Db 231 ValLeuSerProLeuAlaLysAsnLeuPheHisLysAlaIleSerGlnSerGlyValAla 250
|||
QY 879 ATCATCCCTTACCTG--GAGGCCATGATATATGAAAGTAGAGCACTGCAAGTGGT 935
|||
Db 251 LeuThrAlaGlyLeuValLysValLysAsnThrArgProLeuAlaGlnLysIleAlaVal 270
|||
QY 936 GCACATTTCTGTGTGATACATCCGTCAGACCTGAGGCCCTGAGGTGCTGAGGACA 995
|||
Db 271 Ser-----GlyCysLysSerThrThrSerAlaSerMetValHisGlyLeuArgGln 287
|||
QY 996 AAACCTCAAGAGAGCTGCTGACCTGACCCAGAAACAAAGTCTTCACT----- 1046
|||
Db 288 LysThrGlnGlnGlnLeuLeuGlnLysThrLeuLysLysLeuPheSerLeuAsnProLeu 307
|||
QY 1047 -----CGAGTGGTGTGATGCTGCTTCTT 1070
|||
Db 308 HisGlyAspSerArgGlnSerTrpProPheValProThrValLeuAspGlyValValLeu 327
|||
QY 1071 CCTAATGAGCTCTAGATCTATGCTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATC 1130
|||
Db 328 ProLysMetProGlnGlnIleLeuAlaGlnLysAspPheAsnThrValProTrpLysVal 347
|||
QY 1131 GGAGTCAATTAACAGAGTGGTGGCTTCTGCTGCTGCT-----ATGAAGAGAGGCTCTGAGATC 1187
|||
Db 348 GlyIleAsnLysGlnGlnPheGlyTrpIleLeuProThrMetMetAsnTrpProProSer 367
|||
QY 1188 CTCAGTGGCTCCACAGATGCTTGCCTTCCTCATCTGATACAAAC-----ATCTG 1238
|||
Db 368 AspMetLysLeuAspProMetThrAlaThrSerLeuLysLysSerSerPheLeuLeu 387
|||
QY 1239 CACATCCGCTCAGATTTTGCACCTTGTGGCTAATGAATCTTCCATGACAGACATCC 1298
|||
Db 388 AsnLeuProGlnGlnAlaIleProValAlaValGlnLysTrpLeuArgHisThrAspAsp 407
|||
QY 1299 CTGACTGAATCCGAGACGCTTCTGACTGTGAGATGATGTTCTTGTGCTGCT 1358
|||
Db 408 ProAspArgAsnLysAspGlnLeuLeuIleGlyAspValIlePheGlyValPro 427
|||
QY 1359 GCACGTGATCAGAGCTGATATCAGACAGATGCTGGTGCACCTGTCTACTATGAGTTT 1418
|||
Db 428 SerValIleValSerArgGlyHisArgAspAlaGlyAlaGlnThrTyMetTrpGlnPhe 447
|||

QY 1419 CGGACCGGCTTCAGGCTTTGAGACAGACGCGGCTTTTTCGAAAGCCGACCGCT 1478
 Db 448 GlnTrpArgProSerPheSerSerlysmethylpProSerThrValGlyAspHisGly 467
 QY 1479 GATGAATGCCCTTTGTGTGTCGTGTCCTTCGAAAGGCGGACATTTGATGTCGAA 1538
 Db 468 AspGluIleTyrSerValPheGlyValAlaProIleuArgGly----- 481
 QY 1539 GAGCCAGGAGGAGGAGAACTTACTGAGCCGGAAGATGATGAATAACTGGCTACCTTT 1598
 Db 482 GlyThrSerLySGluGluIleAlaAsnSerLysmethylpPheTrpAlaAspPhe 501
 QY 1599 GCTCGAACCGGGAATCTCTAATGGAACGACCTGCTCTGTCGCAAGCTTAACTACT 1658
 Db 502 AlaArgAsnGlyAsnProAsnGlyGlnGlyLeuProHisTrpProGluTyrAspGlnLys 521
 QY 1659 GAGCAGTACCTTCAGCTGACTTGAACATGAGCCTCGACAGACATCAAGAACCGCG 1718
 Db 522 GluGlyTyrLeuGlnIleGlyAlaThrThcGlnGlnAlaGlnLysGlyLysGlu 541
 QY 1719 GTGATTTTGGACACACACCATCCCTGATCCTGTCGCTTCGCAAGCTTCACAGT 1778
 Db 542 ValAlaIleThrPheSerGlu-----LeuLeuAlaMethylpProLeuHisAla 556
 RESULT 15
 S10367
 carboxylesterase (EC 3.1.1.1) ES-10 precursor, microsomal - rat
 N/Alternate names: hydrolyase A
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text_change 09-Jul-2004
 C/Accession: S10367; S12468; S51202; S23460; S14361
 R/Robb, M.; Beaufay, H.; Octave, J.N.
 Biochem. J. 269, 451-458, 1990
 A>Title: Nucleotide sequence of cDNA coding for rat liver pi 6.1 esterase (ES-10), a cat
 A/Reference number: S10367; MUID:90351366; PMID:2386485
 A/Accession: S10367
 A/Molecule type: mRNA
 A/Residues: 1-565 <ROB1>
 A/Cross-references: UNIPROT:Q9R135; EMBL:X51974
 A/Note: 168-61n, 247-Lys, 423-Met, and 506-Asn were also found
 A/Note: the sequence from Fig. 4 is inconsistent with that from Fig. 5 in having 265-Lys
 A/Robb, M.
 submitted to the EMBL Data Library, February 1990
 A/Reference number: S12468
 A/Accession: S12468
 A/Molecule type: mRNA
 A/Residues: 1-264, 'K', 266-565 <ROB2>
 A/Cross-references: EMBL:X51974; NID:956898; PIND:CAA6236.1; PID:956899
 R/Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
 Arch. Biochem. Biophys. 315, 495-512, 1994
 A>Title: Purification and characterization of two rat liver microsomal carboxylesterases
 A/Reference number: S51202; MUID:95077430; PMID:7986098
 A/Accession: S51202
 A/Molecule type: protein
 A/Residues: 19-48 <MOR>
 R/Medda, S.; Proia, R.L.
 Eur. J. Biochem. 206, 801-806, 1992
 A>Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting
 A/Reference number: S23460; MUID:92299008; PMID:1606962
 A/Accession: S23460
 A/Status: Preliminary; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-185, 'Q', 187-422, 'M', 424-505, 'N', 507-565 <MED>
 R/Cross-references: EMBL:X65296; NID:957553; PIND:CAA6391.1; PID:957554
 Biochem. J. 274, 693-697, 1991
 A>Title: Purification and characterization of carboxylesterases from rat lung.
 A/Reference number: S14361; MUID:91190080; PMID:2012599
 A/Accession: S14361
 A/Molecule type: protein
 A/Residues: 19-26, 'D', 28-37 <GAU>
 C/Superfamily: cholinesterase; cholinesterase homology

C/Keyword: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
 F/1-18/Domain: signal sequence #status predicted <SIG>
 F/19-565/Product: carboxylesterase #status predicted <MNT>
 F/50-551/Domain: cholinesterase homology <CHS>
 F/79,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F/221,466/Active site: Ser, His #status predicted
 Alignment Scores:
 Pred. No.: 2,32e-77 Length: 565
 Score: 1121.50 Matches: 245
 Percent Similarity: 56.06% Conservative: 79
 Best Local Similarity: 42.39% Mismatches: 199
 Query Match: 28.96% Gaps: 55
 DB: 2 Indels: 11
 US-10-023-515-1 (1-2158) x S10367 (1-565)
 QY 129 TGGTCTTTTTCGATTCCTCAGCCCTGTTGGACACAGAGTGGGAAATTCGGG 188
 Db 8 TrpLeuPhe-----LeuAlaAlaCythrAlaTrpGlyTyr----- 19
 QY 189 CTTTCTGTAAGGGCCACAGAGAACCGAGCTGGATGATTCAGGCAAGTC 248
 Db 20 ProSerSerPro-----ProValAlaSerThrValLysGlyValLysGlyTyrVal 38
 QY 249 ACTGTCTGGAGACCTGCTGCTGTCGACGTCTTCTCGAGAGTCCCTTGGCTGCC 308
 Db 39 AsnLeuGluGlyPheAlaGlnProValAlaValPheLeuGlyLysProPheAlaLysPro 58
 QY 309 CCGCTGGATCCCGGATTTAAGAACCGGACGCTGCACTGCGCTGGATTAATTGCGCA 368
 Db 59 ProLeuGlySerLeuAlaGlyPheAlaProProGlnProAlaGlnProTrpAsnPheValLys 78
 QY 369 GAAGCACTTCTCACTTAATTTGTCCTTCAGAAC-----TCA 407
 Db 79 AsnThrThrSerTyrProPheMetCysSerGlnAspAlaValGlyGlyGlnValLeuSer 98
 QY 408 GAGTGGCTGCTTTATGATCAACACATGCTCAGCTGATTAACCGAAATTCGAGTGA 467
 Db 99 GluLeuPheThrAsnArgLysGluAlaValPheLeuGlnPhe-----Ser 113
 QY 468 GAAGATGCTCTTACCTGACATCTATGCGCTGCCACCGCATACAGGCTCAAGTC 527
 Db 114 GluAspCysLeuTyrLeuAsnValTyrThrProAlaSerLeuThrLysAsnSerArgLys 133
 QY 528 CCGCTCTGTGTGTGTTCCAGAGGTCCTTCAAGATGCTCAGCTCATCTTGTAT 587
 Db 134 ProValMetValTrpIleHisGlyGlyLysValAlaValAlaThrIleGlnTyrArgLeuGly 153
 QY 588 GGGTCGCGCTGCTGCTATGAGAGCGTGGTGTGCTGCTCAGTACCGCTAGGA 647
 Db 154 GlyGlnValLeuSerAlaHisGlnAsnValValValAlaThrIleGlnTyrArgLeuGly 173
 QY 648 ATTTGTGTTTTCACACATGAGATGAGATGCTCCGGGAATCGGAGCTTCAAGAG 707
 Db 174 IleTrpGlyPhePheSerThrGlyAspGlnHisSerArgLysAsnTrpGlyHisLeuAsp 193
 QY 708 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767
 Db 194 GlnValAlaAlaLeuHisIleTrpValGlnAspAsnIleAlaAsnPheGlyLysAsnProGly 213
 QY 768 TCTGTACCATTTTGGAGAGTCCGCGGACCATTAAGTGTTCAGTCTTACGCTGCT 827
 Db 214 SerValThrIlePheGlyGlnSerAlaGlyLysPheSerValSerAlaLeuValLeuSer 233
 QY 828 CCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTCCATCATCCCT 887
 Db 234 ProLeuAlaLysAsnLeuPheHisArgAlaIleSerGlnSerGlyValValLeuThrSer 253
 QY 888 TACCTGAGGCCCATGATTAATGAGAGAGTGAAGACCTGCGAGTGGTTCGATTTCTGT 947
 Db 254 AlaLeuIleThrThrAspSerLysProIleAlaAsnLeu-----IleAlaThrLeuSer 271

```
OY 948 GGTAACTGCTGAGCTGAGGCCCTGAGAGTGCCTGAGAGCAAAACCTCCAG 1007
    |||
    |||
    |||
Db 272 GlyCysLeuThrThrSerAlaValMetValHisCysLeuArgGlnThrGluSer 291
    |||
    |||
    |||
OY 1008 GAGCTGCTGAGCCCTGAGCAAAACAAAGCTTTC----- 1043
    |||
    |||
    |||
Db 292 GluLeuGluThrSerLeuLysLeuAsnLeuPheLysLeuAspLeuGluLysAsnPro 311
    |||
    |||
    |||
OY 1044 -----ACTGAGGTGTTGATGGTCTTTCCTTCTTAATGAGCCT 1082
    |||
    |||
    |||
Db 312 LysGluSerTyrProPheLeuProThrValIleAspGlyValValLeuProLysThrPro 331
    |||
    |||
    |||
OY 1083 CTAGATCTATTGCTCAGAAACATTAAGCAATTCCTCCATCATCGAGTCAATAAC 1142
    |||
    |||
    |||
Db 332 GluGluIleLeuAlaGluLysSerPheAsnThrValProTyrIleValGlyIleAsnLys 351
    |||
    |||
    |||
OY 1143 CACGAGTGTGCTCTGCTGCTATGAAGAGGCTCTGAGATCCTCAGTGGC----- 1196
    |||
    |||
    |||
Db 352 GlnGluPheGlyTyrIleIleProThrLeuMetGlyTyrProLeuSerGluGlyLysLeu 371
    |||
    |||
    |||
OY 1197 -----TCCAAAGTCCCTTGCCCTCCATGTGATACAAACATCCTGCACATCCG 1247
    |||
    |||
    |||
Db 372 AspGlnLysThrAlaLysSerLeu--LeuTyrLysSerTyrProThrLeuLysIleSer 390
    |||
    |||
    |||
OY 1248 CCTCAGTATTGTCACCTTGCGCTATGATATCTTCATGACAGACACTCCTGACTGAA 1307
    |||
    |||
    |||
Db 391 GluLysMetIleProValAlaIleGluLysTyrPheGlyGlyThrAspAspProAlaLys 410
    |||
    |||
    |||
OY 1308 ATCCGAGACAGCTTCTGAGCTTGTGAGATGTGTTCTTTGTGATCCCTGCACTGATC 1367
    |||
    |||
    |||
Db 411 ArgLysAspLeuPheGlnAspLeuValAlaAspValIlePheGlyValProSerValMet 430
    |||
    |||
    |||
OY 1368 ACAGCTCATATCAGAGATGCTGTGACCTGTCTACTTATGAGTTTCGGACCGC 1427
    |||
    |||
    |||
Db 431 ValSerArgSerHisArgSerPalaGlyAlaProThrPheMetTyrGlnPheGluTyrArg 450
    |||
    |||
    |||
OY 1428 CCTCAGTCTTTGAGACAGACAGCGGCTTTGTCAAGCCGACCGACGCTGATGAATC 1487
    |||
    |||
    |||
Db 451 ProSerPheValSerAlaMetArgProLysThrValIleGlyAspHisGlyAspGluLeu 470
    |||
    |||
    |||
OY 1488 CGCTTGTGTTGGTGGTGCCCTTCTGAAAGGGGACATTGTTATGTTTCAAGAGCCACG 1547
    |||
    |||
    |||
Db 471 PheSerValPheGlySerProPheLeuLys-----AspGlyAlaSer 484
    |||
    |||
    |||
OY 1548 GAGGAGGAGAACTTACTGAGCCGGAAGATGAATAATCGGGCTACCTTGTCTGAACC 1607
    |||
    |||
    |||
Db 485 GlnGlnGluThrAsnLeuSerLysMetValMetLysTyrTyrPalaAsnPheAlaArgAsn 504
    |||
    |||
    |||
OY 1608 GGGAAATCTTAATGGGAACGACTGTCTCTGTGCGCAGTTATATCTGACTGAGCAGTAC 1667
    |||
    |||
    |||
Db 505 GlySerProAsnGlyGlyLeuProHisTyrProGluTyrAspGlnLysGluGlyTyr 524
    |||
    |||
    |||
OY 1668 CTCGAGCTGAGCTTGAACATGAGCCTCGGACAGAGACTCAAGAAACCGCGGGTGAATTT 1727
    |||
    |||
    |||
Db 525 LeuLysIleGlyAlaSerThrGlnAlaIleArgLeuLysAspLysGluValAlaPhe 544
    |||
    |||
    |||
OY 1728 TGGACCAAGACCATCCCTGATCTGTGCTCCGACATGCTCCACAGTCTT 1781
    |||
    |||
    |||
Db 545 TrpSerGlu-----LeuArgAlaLysGluAlaIleGluGluPro 557
    |||
    |||
    |||
```

Search completed: June 15, 2005, 08:55:12
Job time : 72.97 secs

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

this Page blank (uspto)

1	2915	75.3	575	2	Q6NTJ32	Q6nt32	homo sapiens
2	2454	67.3	525	2	Q96DNG	Q96dng	homo sapiens
3	2454	63.4	469	2	Q8NBBC8	Q8nb8	homo sapiens
4	2250	58.1	575	2	Q6AM47	Q6aw7	canis familiaris
5	2094.5	54.1	575	2	Q6AM46	Q6aw6	mus musculus
6	1985	51.3	545	2	Q81034	Q81034	felis silvestris
7	1691.5	43.7	361	2	Q95KH3	Q95kh3	macaca fascicularis
8	1298.5	33.5	571	1	SASB_ANAPL	004791	anas platyrhynchos
9	1282.5	33.1	561	2	Q91W60	Q91w0	mus musculus
10	1281	33.0	559	2	Q8BX48	Q8bx48	mus musculus
11	1276.5	33.0	559	2	Q355J3	Q355j3	meocricetulus auratus
12	1271	32.8	559	2	Q8BM97	Q8bm97	mus musculus
13	1266.5	32.7	556	2	Q6PD87	Q6pd87	mus musculus
14	1263.5	32.6	568	2	Q6CM54	Q6cm54	xenopus laevis
15	1262.5	32.6	561	2	Q70177	Q70177	rattus norvegicus
16	1260.5	32.6	560	2	Q8R097	Q8r097	mus musculus

17	1257.5	32.5	552	2	035535	035535	rattus	norv
18	1239.5	32.0	552	2	069306	069306	xenopus	troc
19	1238	32.0	558	2	08K3R0	08K3R0	rattus	norv
20	1232.5	31.8	558	2	08QQR3	08qqr3	mus	musculu
21	1231.5	31.8	532	1	EST1_RAB1T	P14943	oryzotolagus	
22	1230	31.8	559	1	EST1_HUMAN	000748	homo sapien	
23	1213.5	31.3	551	2	070631	070631	rattus	norv
24	1213.5	31.3	556	2	0766D7	0766d7	felis silve	
25	1213.5	31.3	556	2	0864S9	0864s9	felis silve	
26	1210.5	31.3	557	2	0640T6	0640t6	xenopus	troc
27	1199	31.0	551	1	EST1_MESAU	064149	mesocricetu	
28	1191	30.8	586	2	061PR9	0641k9	homo sapien	
29	1187.5	30.7	568	2	06PG74	06pg74	mus musculu	
30	1187.5	30.7	572	2	08L0S9	08l0s9	mus musculu	
31	1186.5	30.6	554	1	EST1_MOUSE	063880	mus musculu	
32	1183	30.6	555	2	095N05	095n05	canis faml1	
33	1179.5	30.5	571	2	06UWM8	06uwm8	homo sapien	
34	1175	30.4	558	2	07Z6J1	07z6j1	homo sapien	
35	1179	30.3	555	2	08T0Z9	08t0z9	homo sapien	
36	1174.5	30.3	568	2	08V0U1	08v0u1	mus musculu	
37	1169.5	30.2	561	2	091MU0	091mu0	mus musculu	
38	1164	30.1	555	2	097582	097582	sus scrofa	
39	1162.5	30.0	555	2	06GMJ1	06gmj1	brachydanio	
40	1162.5	30.0	555	1	EST1_PIG	029550	sus scrofa	
41	1159.5	29.9	566	2	0464Z1	0464z1	macaca fasci	
42	1150.5	29.7	555	1	EST1_RAB1T	P12337	oryzotolagus	
43	1150	29.7	554	1	EST1_MOUSE	P23953	mus musculu	
44	1149	29.7	555	2	035534	035534	mesocricetu	
45	1144	29.5	549	1	EST1_RAT	P10959	rattus	norv

ALIGNMENTS

RESULT 1			
606NT32			
ID	Q6NT32	PRELIMINARY;	PRT; 575 AA.
AC	Q6NT32;		
DT	05-JUL-2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	FLJ31547 protein.		
GN	Name=FLJ31547;		
OS	Homo sapiens (human).		
OC	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=PCR rescued clones;		
RA	Strasbourg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shemmen C.F., Schutler G.D.,		
RA	Altschuler S.F., Zebberg B., Buetow K.H., Schaefer C.J., Bhat N.K.,		
RA	Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Dickenson L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,		
RA	Brownstein W.J., Ueda T.B., Toshiyuki S., Carrini P., Frange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,		
RA	Boak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,		
RA	Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahy J., Heiton G., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,		
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez C.A., Grimm J., Schmutz J., Myers R.M., Butterfield Y.S.,		
RA	Krzywnicki M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,		
RA	Jones S.J., Maira M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length human		
RT	and mouse cDNA sequences.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=PCR rescued clones;		

RA Director MGC Project;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: BC069501; AAF69501.1; --
DR HSSP: P12337; IKAT.
DR GO: GO:0003824; F: catalytic activity; IEA.
DR InterPro: IPR00379; Ser. ester.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE 575 AA; 63926 MW; 6F5B735BDFC9C09 CRC64;

Alignment Scores:
Pred. No.: 4,51e-206 Length: 575
Score: 2915.00 Matches: 553
Percent Similarity: 97.54% Conservat: 3
Best Local Similarity: 97.02% Mismatches: 6
Query Match: 75.28% Indels: 8
DB: Gaps: 1

US-10-023-515-1 (1-2158) x Q6NT32 (1-575)

QY 129 TGGTGTCTTTCCGATTCACAGCCCTGTTGGGACACAGAGTGGGAAACTGGG 188
DB 14 TTPAlaIleTTPValLeuAlaAlaPro-----ThrylSgLy 25
QY 189 CTTTGTCTGAAGGGCCACAGAGAACACAGAGCTGGAGTGAATTCAGGCAAGCAATG 248
DB 26 ProSerIaGlugLyProGlnArgAsnThrArgLeuGlyTTPrlleGlnGlylSgLyVal 45
QY 249 ACTGTGTGGGAAGCCCTGTGCTGTGAACGTGTCTCGAGTCCCTTTCGCTGCTCC 308
DB 46 ThValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIlePro 65
QY 309 CCGCTGGGATCCCTGGGATTTACGAACCGGACGGCTGCATCCCGCTGGGATTAATTCGGA 368
DB 66 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTTPAspAsnLeuArg 85
QY 369 GAAGCACTCTTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTATGATCA 428
DB 86 GluAlaThrSerTyrProAsnLeuCySLeuGlnAsnSerGlyTTPheLeuLeuAspGln 105
QY 429 CACATGCTGAAGTGATTAACCGAAATTCGAGATGTCAGAAAGCTGCTCTTACCTGAAC 488
DB 106 HisMetLeuLysValHisTyrProLysPheGlyValSerGluAspCySLeuTyrLeuAsn 125
QY 489 ATCTATGCGCTGCCACGCCGATACAGGCTCAGAGTCCCGCTTGTGTGGTTCCTCCA 548
DB 126 IleTyrAlaProAlaHisIleAlaSerThrGlySerIlySLeuProValLeuValTTPPhePro 145
QY 549 GGAGGTGCTTCAACATCGCTCAAGCTTCATCTTGTATGGAGTCCGCTGGCTGCTCAT 608
DB 146 GlyGlyAlaPheLysThrGlySerIaSerIlePheAspGlySerAlaLeuAlaIleTyr 165
QY 609 GAGGACGT 668
DB 166 GluAspValLeuValValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr 185
QY 669 TGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACAGAGTGGCTGCTGCTGG 728
DB 186 TTPAspGlnHisAlaProGlyAsnTTPAlaPheLysAspGlnValAlaAlaLeuSerTTP 205
QY 729 GTCCAGAGAAGATCGAGTTCTTGGTGGGAGACCCGAGCTTGTGACCATCTTTGGGAG 788
DB 206 ValGlnLysAsnIleGlyLysPheGlyGlyAspProSerValThrIlePheGlyGln 225
QY 789 TCCGCGGAGGCATTAAGTGTTCATGCTTATAGTCTTCCCATGGCCAAAGGCTTATTC 848
DB 226 SerIaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLysGlyLeuPhe 245
QY 849 CACAAAGCATTCATGAGAGAGTGGGAGCCATCATCTTCACTGAGAGCCCATGATTAT 908

DB 246 HisIlyAlaIleMetGluSerGlyValAlaIleIleProTyrLeuGluAlaHisAspTyr 265
QY 909 GAGAAGATGAGGACCTTCAGAGGTGTTGTCACATTTCTGTGTACAAATGCTCAGACTCT 968
DB 266 GluLysSerGluAspLeuGlnValAlaHisPheCySgLyAsnAlaSerAspSer 285
QY 969 GAGGCCCTGCTGAGGTGCTGAGCAAAACCCCTCAGAGAGCTTCAGCCCTCAGCAG 1028
DB 286 GluAlaLeuLeuArgCySLeuArgThrLysProSerLysGlyLeuLeuThrLeuSerGln 305
QY 1029 AAAACAAAGCTTTTCACTCGAGTGGTATGATGCTTTCTTCTTAATGAGCTCTGAT 1088
DB 306 LysThrLysSerPheThrArgValAlaAspGlyAlaPhePheProAsnGluProLeuAsp 325
QY 1089 CTATGTCTCGAAGGCAATTTAAAGCAATTCCTTCCATCATCGAGATCAATACCCAG 1148
DB 326 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHisGly 345
QY 1149 TGTGCTTCTGTGCTGCTATGAGAGGCTCTGAGATCTCAGTGGCTCCACAGATCC 1208
DB 346 CySgLyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLysSer 365
QY 1209 CTTGCCCCATCTGATPACAAACATCTCTGACATCCCGCTCAGATTTTGCACCTTTGTG 1268
DB 366 LeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisLeuVal 385
QY 1269 GCTAATGATATCTTCATGACACAGCACTCCCTGACTGAATCCGAGACAGCTTTCGAC 1328
DB 386 AlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeuAsp 405
QY 1329 TTTGCTTGAAGATGTTTCTTTGTGTGCTCTGCACTGATCAACAGCTCGATCAACAGAT 1388
DB 406 LeuLeuGlyAspAlaPhePheValAlaProAlaLeuIleThrAlaArgTyrHisLysGAsp 425
QY 1389 GCTGTGTCACCTGTCTACTCTTATGAGTTTCGGACACGGCTCAGAGCTTGAAGCAAG 1448
DB 426 AlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCySLeuPheGlnAspThr 445
QY 1449 AAGCGGCTTTTGTCAAAGCCGACACGCTGATGAAGTCGCTTGTGTCGGTGGTGGCC 1508
DB 446 LysProAlaPheValIlyAlaAlaPheHisAlaAspGlyValArgPheValPheGlyGlyAla 465
QY 1509 TTCCTGAAGGGGACATTTGTTATGTCGAAGGACCCAGAGAGAGAGCAATTAAGTACGAC 1568
DB 466 PheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuLeuSer 485
QY 1569 CGGAAGATGATGAATAGTGGCTACCTTGTCTCGAACCGGGAATCTTAATGGAAAGCAG 1628
DB 486 ArgLysMetMetLysTyrTTPAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 505
QY 1629 CTGTCTGTGGCGAGCTTATTAATCTGACTGACGAGTACCTTCAGCTGAGTGAATG 1688
DB 506 LeuSerLeuTTPProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMet 525
QY 1689 AGCTTCGACAGAGACTCAAGAACCGCGGGTGAATTTTGGACAGAGACATCCCTCG 1748
DB 526 SerLeuGlyGlnArgLeuLysGluProArgValAspPheTTPHisSerThrIleProLeu 545
QY 1749 ATTCGTGTGCTCTCGACATGTCTCCACAGTCCCTTTTCTTAACCTTCTCTCTCC 1808
DB 546 IleLeuSerIaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 565
QY 1809 CTCGAGCTTTCTTTCTTTTGTGTGCTCT 1838
DB 566 LeuGlnProPhePhePhePheCyAlaIlePro 575

RESULT 2
Q96DN9 PRELIMINARY; PRT; 525 AA.
AC Q96DN9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)


```

Db      61 AspValLeuValValValValGlnTyrArgLeuGlyLeuPheGlyPhePheThrThrTrp 80
QY      672 GATGAGCATGCTCCGGGGAACTGGGGCTTCAAGAGCCAGAGGCTGCTGCTGCTGGGGC 731
Db      81 AspGlnHisAlaProGlyAsnThrPalaPheLysAspGlnValAlaAlaLeuSerTrpVal 100
QY      722 CAGAGAAACATCAGAGTCTTCCGGTGGGAGACCCAGCTGTGAGACCATCTTTGGCGAGTCC 791
Db      101 GlnLysAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGlyGlnSer 120
QY      792 GCGGAGCCCAATAGTGTCTTCAAGCTTATAGTGTCTCCAGCGCAAGCTTATTTCCAC 851
Db      121 AlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaGlyGlyLeuPheHis 140
QY      882 AAGGCAATCAGTGAAGAGTGGGGTGGCCATCATCCCTTACTGAGAGGCCCATATATAGAG 911
Db      141 LysAlaIleIleMetGlnSerIleValAlaIleIleProTyrLeuGlnLysAlaAspTyrGln 160
QY      912 AAGATGAGAGCACTCAGGTGGTGGTGCACATTTCTGTGTAAATGCGTCAGACTCTGAG 971
Db      161 LysSerGlnAspLeuGlnValValAlaHisPheCysGlyLysAsnAsnHisSerAspSerGln 180
QY      972 GCGCTGTGAGGTGCTTGAAGACAAACCCCTTCAAGAGCTGCTGACCTCAGCCAGAAA 1031
Db      181 AlaLeuLeuArgCysLeuAspThrLysProSerLysGlnLeuLeuThrLeuSerGlnLys 200
QY      1032 ACAAAGCTTTTCACTCGAGTGGTGTATGCTGCTTTCTTCCATAGAGGCTTATATCTTA 1091
Db      201 ThrLysSerPheThrArgValValAspGlyAlaAspPheProAsnGlnProLeuAspLeu 220
QY      1092 TTGTCTCAGAAAGCAATTTAAACAATTCCTTCATCATCGAGTCAATAACAGAGTGT 1151
Db      221 LeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnSerHisGlnCys 240
QY      1152 GCGTCTCTGTGCTGCTAATGAGAGGCTCTGTGAGATCTCTGAGTGTCCAAAGATCCCTT 1211
Db      241 GlyPheLeuLeuProMetCysGlyLysAlaProGlnIleLeuSerGlySerAsnLysSerLeu 260
QY      1212 GCGCTCCATCTGATATCAAAACATCCTGCGCATCCCGCTCAGTATTTGACCTTGTGGCT 1271
Db      261 AlaLeuHisAlaLeuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisIleValAla 280
QY      1272 AATGAATCTTCATGACAGAGCACTCCCTGACGTAATCCGAGAGCAAGTCTTGTGACTTG 1331
Db      281 AsnGlnTyrPheHisIleAspLysHisIleSerLeuThrGlnIleArgAspSerLeuLeuAspLeu 300
QY      1332 CTGTGAGATGTGTCTTTGTGTGCTCCCTGCACTGATCAGCTCGATATCAAGAGATGCT 1391
Db      301 LeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAspAla 320
QY      1392 GGTGCACTGTCTTACTTCTATAGAGTTTGGGCAACCGGCTCAGTCTTTGAAGACAGAAAG 1451
Db      321 GlyAlaProValTyrPheTyrGlnPheArgHisArgProGlnCysPheGlnAspThrLys 340
QY      1452 CCGGCTTTTGTGAAGCCGACACAGCGTGAATGAAAGCCGCTTGTGTGGTGGTGGCTTC 1511
Db      341 ProAlaPheValLysAlaAspHisAlaAspGlnValArgPheValPheGlyIleAlaPhe 360
QY      1512 CTGAAGGGGAGACATTTATATGTTTCAAGAGGCAACGAGAGAGAGATTTACTGAGCCGG 1571
Db      361 LeuLysGlyAspIleValIleMetPheGlnGlnGlyAlaThrGlnGlnGlyLysLeuLeuSerArg 380
QY      1572 AAGATGATGAATATCTGGGCTACCTTTGCTCGAACCGGGAATCTTAATGGAAACGACTTG 1631
Db      381 LysMetMetLysTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlnYAsnAspLeu 400
QY      1632 TCTCTGTGGCCGAGCTTATATCTGACTGAGAGTACCTCAGCTGAGACTTGAACATGAGC 1691
Db      401 SerLeuThrProAlaTyrAsnLeuThrGlnGlnTyrLeuGlnLeuAspLeuAsnMetSer 420
QY      1692 CTCGAGACAGAGACTCAAGAAACCGCGGCTGAGATTTTGGACCAAGCACTCCCTGATC 1751
Db      421 LeuGlyGlnArgLeuLysGlnProArgValGlnPheThrThrSerThrIleProLeuIle 440

```

```

QY      1752 CTGTGCTCCGACATGCTCCAGATCCTCTTTCTCTTACTTCTCTCTCTCTC 1811
Db      441 LeuSerAlaAspAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeuLeu 460
QY      1812 CAGCTTTCTTTTCTTTTCTTTTGTGCTCT 1838
Db      461 GlnProPhePhePheCysAlaPro 469

RESULT 4
OS=Am47
ID 06AM47 PRELIMINARY; PRT; 575 AA.
AC 06AM47;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=cauxlin;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN 1;
RP SEQUENCE FROM N.A.
RA Miyazaki M., Yamashita T., Taira H., Suzuki A.;
RT "cauxlin family protein.";
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: belongs to the type-B carboxylesterase/lipase family.
EMBL: AB186392; BAD35015.1; -.
GO GO:0003824; F:catalytic activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR000379; Ser estere.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE 575 AA; 63620 MW; 1005C35B82E1183D CRC64;

Alignment Scores:
Pred. No.: 4,93e-157 Length: 575
Score: 2250.00 Matches: 417
Percent Similarity: 86.75% Conservative: 61
Best Local Similarity: 75.66% Mismatches: 73
Query Match: 58.11% Indels: 0
Gaps: 0

US-10-023-515-1 (1-2158) x 06AM47 (1-575)
QY      186 GGGCTTCTGCTGAAGGGCCACAGAGAAACACAGCTGGATGATTCAAGGCAAGCA 245
Db      25 GlyProAlaAlaAspAlaProValArgSerThrArgLeuGlyTrpValArgGlyLysGln 44
QY      246 GTCACTGTGCTGGAGAGCCGTGTGCTTGAACGTGTCTCTGAGTCCCTTGTGCTGCT 305
Db      45 AlaThrValLeuGlySerThrMetProValAsnValPheLeuGlyIleProPheAlaAla 64
QY      306 CCCCCGCTGGAGTCCCTCGATTTACGAACCCGAGCTGTCATGCCCTCGGATAACTTG 365
Db      65 ProProLeuGlyProLeuAlaArgPheLysArgProLysProAlaLeuLeuThrAsnAspSer 84
QY      366 CGAGAAAGCACCTCTTACCTTAATTTGTGCTCCGAACTCAGAGTGGTGTCTTGAAT 425
Db      85 ArgAspAlaThrSerTyrProLysLeuGlyValSerGlnAsnSerValTrpLeuLeuSerAsp 104
QY      426 CAACATGCTCAAGTCAATTAACCCGAATTCGAGAGTGCAGAAAGTGCCTTACTTG 485
Db      105 GlnHisPheLeuLysValHisTyrProAsnLeuGlnValSerGlnAspCysLeuTyrLeu 124
QY      486 AACATCTATGGCTGCCCCAGCCGATACAGGCTCCAGGCTCCCGTCTTGTGTGCTTC 545
Db      125 AsnIleTyrAlaProAlaHisAlaAsnThrGlySerLysLeuProValMetValTrpPhe 144
QY      546 CAGAGAGTGCCTTCAAGACTGCTCAGCTTCATCTTGAATGGGTGCGCTGCTGCC 605

```

Db	145	ProGlyValAlaPheIleuThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla	164
Qy	606	TATGAGAGCTGTGCTGTGTGCTCTCCAGATACCGGCTAGGAATATTTGGTTCTTACC	665
Db	165	TyrGluAspValLeuIleValIleThrGlnTyrIleGluGlyIlePheGlyPhePheLeu	184
Qy	666	ACATGGGATTCAGCAAGCTCCCGGGAACTGGGCTTCAAGACCAAGTGGCTGTGCC	725
Db	185	ThrGlyAspGlnIleValAlaProGlyAsnTyrAlaPheLeuAspGlnLeuAlaAlaLeuThr	204
Qy	726	TGGGCTCAGAAACATGAGATTCTTCGCTGGGAGCCCGAGCTGTGACATCTTTGGC	785
Db	205	TyrValIleGlnIleuAsnIleGluPhePheGlyIleAspProIleSerValIleThrIlePheGly	224
Qy	786	GAGTCCCGGGAGCCATTAAGTGTTCATGCTTATCTGCTCCGATGGCCAAAGCTTAA	845
Db	225	GlnSerAlaGlyAlaIleSerValSerGlyLeuValLeuSerProMetAlaSerGlyLeu	244
Qy	846	TTCCACAAAGCCATCAGAGAGGGGGGTGGCCATCTCCCTTACCTGAGGCCATGAT	905
Db	245	PheIleValAlaIleMetGluSerGlyValAlaIleIleProIleLeuIleArgAlaProAsp	264
Qy	906	TATGAGAAAGTGAAGACTTCGACAGTGGTTGCACATTTCTGTGTAAACAATGGCTGAC	965
Db	265	AspIleIleArgAsnIleuPheIleuGlnValIleAlaArgIleCysAspIleCysAsnIleSerAsp	284
Qy	966	TCTGAGGCCCTGCTGAGAGTCCGAGAGCAAAACCTCCAGAGAGCTGTGACCTCAGC	1025
Db	285	SerValAlaIleuLeuGlnCysLeuAlaArgAlaIleSerSerGlnIleuLeuAspIleAsn	304
Qy	1026	CAGAAAACAAGTCTTTCACCTCGAGTGTGATGTGCTTTCTTCTTATAGAGCTCTTA	1085
Db	305	LysAlaThrIleYsSerPheThrArgValValAspGlyPhePhePheAspIleProIleu	324
Qy	1086	GATCTATTGCTCAGAAAGCATTTAAAGCAATCTCTCATATGAGATGCAATTAACCA	1145
Db	325	AspLeuLeuThrGlnIleuYsThrPheAsnSerIleProSerValIleGlyValAlaAsnIle	344
Qy	1146	GAGTGTGCTCTCTGCTGCTATGAGAGAGCTCTCGAATCTCAGTGGGCTCCAAAG	1205
Db	345	GluCysGlyPheLeuLeuProMetIleuGlnPheProGlnIleLeuGlyIleSerAsnIle	364
Qy	1206	TCCCTTGGCCCTCATTGTGATACAAACATCTGCAATCCGCGCTCAGTATTGCACTT	1265
Db	365	SerIleAlaLeuIleIleuIleIleIleValIleuIleIleProAsnGlnTyrLeuTyrIleu	384
Qy	1266	GTGGCTAATGAATACCTTCATGACAAGACCTCCCTGATGAAATCCGAGACGTTCTTG	1325
Db	385	ValAlaAspGlnTyrPheTyrAsnIleYsSerProValIleuIleAspSerPheLeu	404
Qy	1326	GACTTGCTTGAGAGTGTGTTCTTTGTGCTCCCTGACATGACAGCTCGATATCAGAGA	1385
Db	405	AspLeuLeuGlyAspValPhePheValValIleProGlyValIleThrAlaArgTyrIleAsp	424
Qy	1386	GATGCTGGAGCACTGTACTCTATAGATTTGGGACCGGCGCTCAGTCTTTGAAGAC	1445
Db	425	AspAlaGlyAlaProValTyrPheTyrGlnPheGlnIleAspProGlnIleCysLeuYsAsp	444
Qy	1446	ACGAAAGCCGCTTTGTTCAAAGCCGACCGATGAGAGCTCGCTTGTGTGCTGGGT	1505
Db	445	ThrArgProAlaPheValIleValIleAspIleSerAspArgIleIleArgPheValIleGlyIle	464
Qy	1506	GCTTTCCTGAGAGGGGACATGTTATGTGTGAGAGAGACCCGAGAGAGGAAGTTACTG	1565
Db	465	AlaPheLeuYsGlyAsnIleValIleMetPheGlnGlyAlaThrGlnGlnIleuYsLeuLeu	484
Qy	1566	AGCCGAAAGATGATGAATACTGGGCTACCTTTGCTGAAACCGGAGATCTATGGGAAC	1625
Db	485	SerArgIleMetMetArgTyrTyrAlaAsnPheAlaArgThrIleYsAspProAsnGlyIleu	504
Qy	1626	GACCTGCTCGTGGCCAGCTTATATCTGACAGAGCACTCCAGCTGAGCTTGAAC	1685

[illegible]

[illegible]

```

Db      466 PheLeuYlaG1yAspAlaValMetPheGluGlnAlaThrTrgUlnGluIlybLeuSer 465
Oy      1569 CGAAGAATGATGAATAATCTGGGGCTACCTTTGCTGCAACCGGAAATCTTAATGGGAACGAC 1623
Db      486 ArgIlyMetSerIlyThrTrpAlaAsnPheAlaArgSerGlyAspProAsnGlyAlaAsp 505
Oy      1629 CTGTCTCTGTGGCCACGCTTAATATCTGACTGAGCGATCTCTCCACGCTGAGATTGAACATG 1689
Db      506 LeuProProTrpProValIlyTrpAspGluAsnGluTrpIleuGluIleuAspValAsnIle 523
Oy      1689 AGCCTGGACAGAGACTCAAAAGAACCGGGGGATTTTGGACACAGACCATCCCTCGT 1749
Db      526 SerThrGlyArgTrpArgLeuIlyAspGlnArgAlaGluPheTrpThrAspTrnLeuProleu 545
Oy      1749 ATCCTGTGCGCCCTCCGACATGCTCCACAGTCCCTCTTTCTTCTTAATCTTCTCTCTC 1809
Db      546 IleuSerIleuSerIleuSerIleuSerProThrPheSerIleuIleuLeuSerIleu 565
Oy      1809 CTCACGCTTTCTTTTCTTTTCTTTGCT 1835
Db      566 LeuSerProValIleuSerIleuAla 574

RESULT 6
ID      081034      PRELIMINARY;      PRT;      545 AA.
AC      081034;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Carboxylesterase-like urinary excreted protein.
GN      Name=Cauxin;
OS      Fells silvestris catus (Cat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Fells.
OX      NCBI_TaxId=9685;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      MEDLINE=2458314; PubMed=12401131; DOI=10.1042/BJ20021446;
RA      Miyazaki M., Kamie K., Soeta S., Taira H., Yamashita T.;
RT      "Molecular cloning and characterization of a novel carboxylesterase-
RT      like protein that is physiologically present at high concentrations in
RT      the urine of domestic cats (Fells catus).";
RC      Blochem. J. 370:101-110(2003).
CL      1-1 SMILIARY: Belongs to the type-B carboxylesterase/lipase family.
DR      EMBL; AB045377; BAC25277.1; -.
DR      HSP; P12337; 1k4Y.
DR      GO; GO:0016787; F:hydrolase activity; IEA.
DR      InterPro; IPR002018; CarboxesteraseB.
DR      InterPro; IPR000379; Ser_estr.
DR      Pfam; PF00135; Coesterase; 1.
DR      PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR      PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW      Hydrolase.
SQ      SEQUENCE 545 AA; 60505 MW; 9F73FA693D271FA9 CRC64;

Alignment Scores:
Pred. No.:      1 68e-137      Length:      545
Percent:      1985.00      Matches:      373
Percent Similarity:      83.27%      Conservative:      60
Best local Similarity:      71.73%      Mismatches:      85
Query Match:      51.27%      Indels:      2
DB:      2      Gaps:      2

US-10-023-515-1 (1-2158) x 081034 (1-545)
Oy      186 GGGCGCTTCTGCTGAAGCGGCACAGAGAACACAGCGCTGGAGTGAATCAAGGCAACAA 245
Db      25 GlyProAlaAlaAspAlaProValAlaArgSerThrArgLeuGlyTrpValArgGlyIlyeGln 44
Oy      246 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGATTCCTCGGAGTCCCTTTGCTGCT 305

```

```

Db      45 ThrThrValLeuGlySerThrValProValaAsnMetPheLeuGlyIleProTyrAlaIa 64
Qy      306 CCCCGCTGGGATCCCTGGGATTTAGAACCCGCGAGCTGCATCCCTCGGATTAACTTG 365
Db      65 ProPheLeuGlyProLeuArgPheLeuGlnProLysProAlaLeuProGlyAsnAspPhe 84
Qy      366 CGAAGAGCCACTCTCACTTAATTTGGCTCCGACAGCTGAGTGGCTGCTCTTGAT 425
Db      85 ArgAsnAlaThrSerTyrProLysLeuCybPheGlnAspLeuGlnTyrPheValaSerTyr 104
Qy      426 CAACGATGCTCAAGTGTGATTAACCGAAATTCGAGTGTCAAGAACTGCTTACCTG 485
Db      105 GlnHisValLeuValArgTyrProLysLeuGlnAlaSerGlnAspCysLeuTyrLeu 124
Qy      486 AACATCTATGGCGCTCCGACGCGATACAGGCTCAAGCTCCCGCTTGGTGTGCTTC 545
Db      125 AsnIleTyrAlaProAlaHisAlaAspAsnGlySerAsnLeuProValaMetValTyrPhe 144
Qy      546 CAGAGAGTGCCTTCAGAGCTGGCTCAAGCTCTTGTATGGGTCCGCTGGCTGGC 605
Db      145 ProGlyValaPheLysMetGlySerAlaSerSerPheAspGlySerAlaLeuAlaIa 164
Qy      606 TATGAGAGCTGCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665
Db      165 TyrGlnAspValaLeuIleValThrThrGlnTyrArgLeuGlyIlePheGlyPheAsp 184
Qy      666 ACATGGGATCAGCATGCTCCGGGAGACTGGGCTTCAGAGACAGGCTGGCTGTCTCC 725
Db      185 ThrGlyAspGlnHisAlaArgGlyAsnTyrAlaLeuLeuAspGlnValaAlaLeuThr 204
Qy      726 TGGGTCAGAGAAACATCGATCTTCCTGGTGGGAGACCCGAGCTGTGACCATCTTGGC 785
Db      205 TrpValaArgAspAsnIleGlnPhePheGlyAspProArgSerAlaThrIlePheGly 224
Qy      786 GAGTCCGCGGAGCCATAGTCTTTTCTAGCTTATCTGTCTGTCCCAAGCAAGCTTA 845
Db      225 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAsnGlyLeu 244
Qy      846 TTCACAAAGCCATCATGAGAGTGGGCTGGCCATCATCCCTTAACCTG--GAGGCCAT 902
Db      245 PheHisValaAlaIleMetGlnSerGlyValAlaIleLeuProLeuMetArgProPro 264
Qy      903 GATTATGAGAGAGTGAAGACTGCTGAGTGTGTGACATTTCTGTGTGAACATCCGCA 962
Db      265 GlyAspGlnArgLysValAspLeuGlnValIleAlaHisGlyIleCysHisAlaSer 284
Qy      963 GACTCGAGGCTGCTGAGTGTGTGAGTGTGTGAGACAAACCTCCAGAGAGTGTGACCTC 1022
Db      285 AspSerAlaAlaLeuLeuGlnCysLeuArgAlaLysProSerGlnGlnLeuMetAspIle 304
Qy      1023 ACCGAGAAACAAAGCTTTCACTCGAGTGTGATGTGCTTTCTTCTTAATGAGCTT 1082
Db      305 SerLysLysLeuThrPheSerIleProValIleAspAspPhePheProAspLysPro 324
Qy      1083 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATAC 1142
Db      325 ValAlaLeuLeuThrGlnLysAlaPheAsnSerValProSerIleIleGlyValaAsn 344
Qy      1143 CACGAGTGTGCTGCTGCTGTGCTTGAAGAGAGCTCTGAGATCTCTCAAGTGTCCAAC 1202
Db      345 HisGlnCysAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlyIleSerAsn 363
Qy      1203 AAGTCCCTTGGCTTCATCTGATCAAAACATCCCTGACATCCGCGCTCAGTATTGGAC 1262
Db      364 ArgSerLeuAlaLeuTyrLeuValHisThrPheLeuAsnIleProThrGlnTyrLeuHis 383
Qy      1263 CTGTGGCTAATGAAATCTTCATGACAGACATCTCCGACTGAATCCGAGACAGTCTT 1322
Db      384 LeuValAlaAspHisTyrPheTyrAsnLysHisSerProValGlnIleArgAspSerPhe 403
Qy      1323 CTGAGCTTGTGAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1382
Db      404 LeuAspLeuLeuGlnLysAspValLeuPheValValProGlyValaValThrAlaArgTyrHis 423

```

```

Qy      1383 AGAGATGCTGTGACACTGTGTCTTATTAAGTGTTCGGACCGGCTCAAGTCTTTGAA 1442
Db      424 ArgAspAlaGlyAlaProValTyrPheTyrGlnPheGlnHisProProGlnCysLeuAsn 443
Qy      1443 GACACGAGCCGAGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCGT 1502
Db      444 AspThrArgProAlaPheValaLysAlaAspHisSerAspGlnIleArgPheValaPheGly 463
Qy      1503 GGTGCTTCTCAAGGGGAGACATTTGTTATTTTCGAAGAGCCAGAGAGAGAAAGTTA 1562
Db      464 GlyAlaPheLeuLysGlyAspIleValaMetPheGlnGlyAlaThrGlnGlnGlyLeu 483
Qy      1563 CTGACCCGAGAGATGATGAATTAATCTGAGTACTTGTGTCAACCGGGAATCTTAAGG 1622
Db      484 LeuSerArgLysMetMetCysGlyTyrTrpAlaAsnHisAlaGlnThrGlyAspProAsnGly 503
Qy      1623 AACGACCTGTCTGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1682
Db      504 GlnGlyValProLeuTyrProAlaTyrThrGlnSerGlnGlnTyrLeuLysLeuAspLeu 523
Qy      1683 AACATGACCTCGGACAGAGACTCAAGAAACCGGAGTATTTTGGACACAGCATC 1742
Db      524 SerValSerValaGlyGlnLysLeuLysGlnGlnValaGlnPheThrMetAsnThrIle 543

RESULT 7
Q95KH3 PRELIMINARY; PRT; 361 AA.
AC Q95KH3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey) .
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euleleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_Taxid=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Temporal lobe right;
RA Oesda N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
DR EMBL: AB060873; BAB46884.1; -.
DR HSPB: P12337; 1K4Y.
DR InterPro: IPR002018; CarboxylesteraseB.
DR Pfam: PF00135; Coesterase; 1.
KW Hypothetical protein.
SQ SEQUENCE 361 AA; 41032 MW; AC62AB6E78C5F1A1 CRC64;

Alignment Scores:
Pred. No.: 6,37e-116 Length: 361
Score: 1691.50 Matches: 327
Percent Similarity: 96.23% Conservatave: 5
Best Local Similarity: 94.78% Mismatches: 12
Query Match: 43.69% Indels: 1
DB: 2 Gaps: 1

US-10-023-515-1 (1-2158) x Q95KH3 (1-361)
Qy      819 ATATGCTTCCCATGGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGTGCC 878
Db      17 IleLeuSerProMetAlaGlnGlyLeuPheHisValaIleMetGlnSerGlyValaIa 36
Qy      879 ATCATCCCTTACCTGGAGAGCCCATGATTTATGAAGAGTGAAGACTTGAAGTGTGCA 938
Db      37 IleIleProTyrLeuLysAlaTyrAspTyrGlnLysSerGlnAspLeuGlnValaIaIa 56
Qy      939 CATTTCTGTGTAACAATGCTCAGACTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGACAAA 998
Db      57 HisPheCysGlnLysAsnAlaAspSerSerCysLeuLeuAsnArgCysLeuArgTyrLys 76

```

```

OY 999 CCTCCAGAGAGCTGCTGACCTGACCCGAGAAACAAAGCTTTTCACTGAGTGTTGAT 1058
DB 77 SerSerlySGluLeuLeuThreuserglnlysalalysSerPheThrArgValValAsp 96
OY 1059 GGTGCTTTCTTTCTTAATGAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAAT 1118
DB 97 GlnAlaPhePheProAsnGlnProLeuGlnLeuLeuSerGlnlysalalalyalaile 116
OY 1119 CCTTCATCATCGAGAGTCAATTAACCAAGTGTGGCTTCTGCTGCTCTATGAGAGAGCT 1178
DB 117 ProSerIleIleGlyValAlaAsnAsnHISGluCySGlyPheLeuLeuProMet---GluAla 135
OY 1179 CCTGAGATCTCTGAGGCTCCGAAAGTCCCTTGGCCCTCATCTGATTAACAAACATCTCTG 1238
DB 136 ProGlnIleLeuSerGlySerAsnlySerLeuAlaLeuHISleuIleGlnAsnIleLeu 155
OY 1239 CACATCCCGCTCAGATATTGACCTTGCTGCTAAATGAACTTGCATGACAGCACTCC 1298
DB 156 HisIleProProGlnlyrleuHISleuValAlaAsnGlnlyrPheHISAspIySHISser 175
OY 1299 CTGATCGAAATCCGAGACAGTCTTCTGACCTTGTGAGATGTGTCTTTGTGCTCCT 1358
DB 176 LeuThrGlnIleArgAspGlyLeuLeuAspLeuLeuGlnlyAspValPhePheValValPro 195
OY 1359 GCACGTATCACAGCTCGATATACAGAGATGCTGGGACCTGCTGCTACTTCTATGAGTTT 1418
DB 196 AlaLeuIleHISAlaArgIyHISArgAspAlaGlyAlaProValIyPheIyroluPhe 215
OY 1419 CGGACCGGCTCAGTGTCTTGTGAGACAGAACCGGCTTTTGTCAAGCCGACCACT 1478
DB 216 ArgHISArgProGlnCyPheGlnAspThrIySProAlaPheValIyAlaAspHISAla 235
OY 1479 GATGAAGTCCGCTTGTGTGCTGCTGCTGCTCTTCTGAAAGGAGCAATGTTATGTTGAA 1538
DB 236 AspGlnValArgPheValPheGlyGlyAlaPheLeuLySGlyAspIleValMetPheGln 255
OY 1539 GGAGCGACGAGGAGAGAGAGTACTGAGCCGGAAGATGAAATPACTGGGCTACCTT 1598
DB 256 GlnAlaThrGlnGlnGlnlyLeuLeuSerArgIySMeMetLyIyTrpAlaThrPhe 275
OY 1599 GCTCGAACCGGAGATCTTAATGGAACGACTGCTCTGTGCGCAGCTTATATCTGACT 1658
DB 276 AlaArgThrGlyAsnProAsnGlnlyAspAspLeuProLeuTrpProAlaIyTrsMetLeuThr 295
OY 1659 GAGCAGTACTCTCAGCTGACTTGAACATGAGCTCTGACAGACAGACTCAAGACCGCG 1718
DB 296 GlnGlnIyIyLeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLySGlnProArg 315
OY 1719 GTGATTTTGTGACACGACCACTCCCTGATCTCTGCTCCGACATGCTCCACAGT 1778
DB 316 ValGlnPheTrpThrIlePheProLeuIleLeuSerAlaSerAspThrIleuHISser 335
OY 1779 CCTCTTTCTCTTAACTTCTCTCTCTCTCCTGAGCTTTCTTTTCTTTTGTGCTCCT 1838
DB 336 ProLeuSerPheLeuIlePheLeuSerLeuLeuGlnProPhePheSerCySValPro 355
OY 1839 TGAGAGATTATCTTT 1853
DB 356 TrpGlnValIlePhe 360

```

```

RESULT 8
SASB_ANAPL STANDARD; PRT; 557 AA.
AC Q04791;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DB Pctly acyl-CoA hydrolase precursor, medium chain (EC 3.1.2.-)
DB (Thioesterase B).
DB Anas platyrhynchos (Domestic duck).
DB Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.

```

```

OX NCBI_TaxID=8839;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-65.
RC TISSUE=utrocytic gland;
RX MEDLINE=9310623; PubMed=8314791;
RA Hwang C.-S., Kolatukudy P.E.;
RT "Molecular cloning and sequencing of thioesterase B cDNA and
RT stimulation of expression of the thioesterase B gene associated with
RT hormonal induction of peroxisome proliferation.";
RL J. Biol. Chem. 268:14278-14284(1993).
CC -1- FUNCTION: Fatty acid biosynthesis chain termination and release of
CC the free fatty acid product is achieved by hydrolysis of the thio
CC ester by a thioesterase. This thioesterase may be associated with
CC peroxisome proliferation and may play a role in the production of
CC 3-hydroxy fatty acid diester phenomones.
CC -1- TISSUE SPECIFICITY: Highest levels in utrocytic gland, much lower
CC in liver and kidney.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC
DR EMBL; L05493; AAA49223.1; -.
DR PIR; A47162; A47162.
DR HSP; 07540; 1K4Y.
DR InterPro; IPR002018; CarboxylesteraseB.
DR InterPro; IPR00379; Ser estere.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
DR KW Direct protein sequencing; Fatty acid biosynthesis; Hydrolyase; Signal.
FT SIGNAL 1 25
FT CHAIN 26 557 Fatty acyl-CoA hydrolase.
FT ACT_SITE 227 227 Acyl-ester intermediate (By similarity).
FT ACT_SITE 345 345 Charge relay system (By similarity).
FT ACT_SITE 460 460 Charge relay system (By similarity).
FT DISULFID 93 122 By similarity.
FT CARBOHYD 476 476 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 557 AA; 61637 MW; 03E35D9DA037F6B0 CRC64;

Alignment Scores:
Pred. No.: 7.14e-87 Length: 557
Score: 1298.50 Matches: 265
Percent Similarity: 62.77% Conservative: 79
Best Local Similarity: 48.36% Mismatches: 177
Query Match: 33.54% Indels: 27
Gaps: 8

US-10-023-515-1 (1-2158) x SASB_ANAPL (1-557)
OY 183 ACTGGCCCTTCTGCTGAGCGCCACAGAGAAACAGCTGCGATTCAGGGCAAG 242
DB 23 ThrGlnGlnlysalalaglnProGlnValAlaThrAsnIyrglySerValArgGlyIy 42
OY 243 CAACTGCTGCTGCTGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
DB 43 GlnVallysalalalalalalaglnArgSerValanValPheLeuGlyLeuProPheAla 62
OY 303 GCTCCCGCTGAGATCCCTGCGATTATGAAACCCGACGCTGATCGCCCTGGATAC 362
DB 63 LysProProValGlyProLeuArgPheSerGlnProGlnProGlnProGlnProGlnly 82
OY 363 TTGCGAGAGACACTCTTACCTTAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
DB 83 ValArgAspAlaAlaAlaSerIyProPheMetCySLeuGlnAspLyS-----ValLeu 99
OY 423 GATCAACATGCTC-----AAGGTGATTACCCGAAATTCGAGGTGCA 467

```

```

Db      100 GlyInTyLeuSerAspAlaIleThrAsnArgLySGluYValArgLeuGlnIleSer 119
Qy      468 GAAGATGCTCTTACCTGAACATCTATGGCGCTGCCCGCATACAGGCTCCAGATC 527
Db      120 GlnuPcybLeuTyLeuAsnValTyThrProValSerThrgluGlnGlnuYbLeu 139
Qy      528 CCCGCTTGGTGGTCCCGAGAGGCTCTTCAAGACCTGCGCTCAGCCTCCATCTTGAT 587
Db      140 ProValPheValTrpIleHieGlyGlyLeuValSerGlyAlaAlaIleSerTyrAsp 159
Qy      588 GGGTCCGCGCTGGCTATGAGAGACCTGCTGTGTGGTCGTCAGTACCGGCTAGA 647
Db      160 GlySerAlaLeuAlaAlaPheAspAsnValValValAlaThrIleGlnTyArgLeuGly 179
Qy      648 ATATTGGTTCTTCCACCATGGGATCAGCATGCTCCGGGAACTGGGCTTCAGAGAC 707
Db      180 IleAlaGlyTyRPheserThrGlyAspLyehiAlaArgGlyAsnTrpGlyTyLeuAsp 199
Qy      708 CAGGTGGTCTCTCTCTGCTGGGTCAGAGAACATTCAGTTCTTCCGTGGGACCCGAC 767
Db      200 GlnValAlaAlaLeuGlnTrpIleGlnGlnuAnIleIleHiePheArgLyAspProGly 219
Qy      768 TCTGTGACCATCTTTGGCAGTCCGCGGAGCCATTAAGTTTCTAGTCTTAACTGCT 827
Db      220 SerValThrIlePheGlyGlySerAlaGlyValSerValSerAlaLeuValLeuSer 239
Qy      828 CCCATGGCCAAAGGCTTATTCACAAAGCCATCATGGAGTGGGCTGCATC--ATC 884
Db      240 ProLeuAlaYbGlyLeuPheHieLyAlaIleSerGlnuSerGlyThrAlaValArgIle 259
Qy      885 CTTTACCTGGAGGCCATGATTGAGAAGTGAAGACCTGCAAGTGGTGTGCATTTTC 944
Db      260 LeuPheThrGlu-----GlnProGlnGlnGlnAlaGlnArgIleAlaAla 275
Qy      945 TGTGTAAACATGCTGAGACTGAGGCTGAGGCTGAGGCTGAGGAGCAAAACCTCC 1004
Db      276 AlaGlyCybGlyuYbSerSerSerAlaAlaLeuValGlnCybLeuArgGlyuYbThrglu 295
Qy      1005 AAGGACCTG-----CTGACCCCTCAGCCAGCAAAACAAAGTCTTTTCACTGAGTGTGAT 1058
Db      296 AlaGlnMeGlnGlnIleThrLeuYbMetProPromePheIleSerAlaSerLeuAsp 315
Qy      1059 GGTGCTTTTCTTCTTATGAGCTCTTATGATCTATTGTCAGAAAGCATTTAAACAT 1118
Db      316 GlyValPhePheProYbSerProArgGlnLeuLeuSerGlnuYbValIleAsnAlaVal 335
Qy      1119 CTTTCATATATGGAGTCAATTAACACAGAGTGGCTTCTGCTGCT--ATGAAAGAG 1175
Db      336 ProTyRleIleGlyValaAsnAsnCysGlnuPheGlyTrpIleLeuProArgMetLeuYb 355
Qy      1176 GCTCTGAGATCTCTGAGTCCCAACAGTCCCTTGGCTCATCTATCAAAACATC 1235
Db      356 PheProGlnuPheThrGlnGlyLeuGlnuYbAspValAlaArgGlnValLeuGlnIleSerThr 375
Qy      1236 CTGCACATC-----CCGCTCAGATTGTCACCTTGTGGTAAATGAATAC 1280
Db      376 LeuAlaLeuSerPheLyGlyAlaProSerAspIleValaAspLeuValTyAsnGlnuYb 395
Qy      1281 TTCATGACAGACATCCCTGATCTAAATCCGAGACAGTCTTTGAGACTTCTTGAGAT 1340
Db      396 IleGlyValAlaGlnuAsnArgAlaGlnValArgAspGlyLeuLeuAspSerIleAlaAsp 415
Qy      1341 GTGTTCTTTGTGCTCCCTGACATGTCACAGTCACTGATATGCGTGGGCACT 1400
Db      416 ProLeuPheValPheSerAlaValaGlnuValaAlaArgHieHieAspAspAlaGlyAsnPro 435
Qy      1401 GTTACTTCTATAGTTTGGCACCAGCTGCTGAGTGTGGAAGACAGAACCGGCTTT 1460
Db      436 ValTyRPhetyRglnuPheGlnHieArgProSerSerAlaAlaGlyValaProGlnuPhe 455
Qy      1461 GTCAAAGCCGACGAGTGAAGTCCGCTTGTGTTGCTGCTGCTGCTGAGGG 1520
Db      456 ValYbAlaAspHieAlaAspGlnIleAlaPheValPheGlyuYbProPheLeuAlaGly 475

```

```

Qy      1521 GACATTGTTATGTTCCGAAGACCAAGAGAGAGAAATTATCTGAGCCGGAAGATGATG 1580
Db      476 Asn-----AlaThrGlnGlnGlnAlaYbLeuSerArgThrValMet 489
Qy      1581 AAATATCGGCTACCTTTTCTCTGAAACCGGAAATCTTAATGGGAACGACTGCTCTG 1640
Db      490 LybTyRTrpThrAsnPheAlaArgAsnGlyAsnProAsnGlyGlnGlyLeuValHieTrp 509
Qy      1641 CCAGCTTATATCTGACATGATCCGACGATCCAGCTGAGACTTGAACATGAGCTCGACAG 1700
Db      510 ProGlnTyRAspMeAspGlnuArgTyLeuGlnuIleAspLeuThrglnuYbAlaAlaYb 529
Qy      1701 AAGCTCAAGAAACCGCGGTGAGATTTTTGAACACGACCATCCCTGATCTGCTGACC 1760
Db      530 LybLeuYbGlnuArgLybMeGlnuPheTrpMetGlnuThrglnuIleMeSerAsp 549
Qy      1761 TCCGACATGCTCCACAGTCTCTT 1784
Db      550 ArgArgArgLybHieThrAspLeu 557

RESULT 9
ID      Q91MG0      PRELIMINARY,      PRT,      561 AA.
AC      Q91MG0;
DT      01-DEC-2001 (TrEMBLrel. 19, Created)
DT      01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT      25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE      Carboxylesterase 2 (Carboxylesterase MJ3 precursor).
GN      Name=Ces2; Synonyms=ces2a3;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Colon, and Kidney;
RX      MEDLINE=22388257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA      Strauberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA      Klausner R.D.; Collins F.S.; Wagner L.; Shennan C.M.; Schuler G.D.;
RA      Altschul S.F.; Zeeberg B.; Buettow K.H.; Schaefer C.F.; Hsieh F.;
RA      Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA      Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA      Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA      Brownstein M.J.; Uedlin T.B.; Toshynski S.; Carninci P.; Prange C.;
RA      Bock S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Mullahy S.J.;
RA      Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;
RA      Villalon D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA      Fahy J.; Helton E.; Kettelman M.; Madan A.; Rodriguez S.; Sanchez A.;
RA      Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA      Blakeley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA      Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.W.; Butcherfield Y.S.;
RA      Krzyzaniak M.I.; Skalska U.; Smalins D.E.; Scherch A.; Schein J.E.;
RA      Jones S.U.; Maizra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Kidney;
RA      Strauberg R.;
RN      Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Colon;
RA      Strauberg R.;
RN      Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
[4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Colon;
RA      Strauberg R.;
RN      Submitted (Jun-2002) to the EMBL/GenBank/DBJ databases.

```


Alignment Scores:

Pred. No.: 3e-85 Length: 559
 Score: 1276.50 Matches: 262
 Percent Similarity: 64.58 Conservative: 79
 Best Local Similarity: 49.628 Mismatch: 172
 Query Match: 32.974 Indels: 15
 DB: 2 Gaps: 9

US-10-023-515-1 (1-2158) x 035533 (1-5559)

```

QY 177 GGAAGAACTGGGCTTTCTGCTGAGAGGCGCAGAGAAACACCGAGCTGGAGTTCAG 236
DB 26 GYGINAAPSERProSer-----ProIIeaGAnThrH1eThnGlyGlnValArg 42
QY 237 GGCAGCAAGCTCATCTGTCTGGAGAGCCCTGAGCTGGAGACGTTCTCTGGAGTCCC 296
DB 43 GYAGLeuValH1eValLlyAaPThrGlnValAaPValH1eThnPheLeuGlyLePro 62
QY 297 TTGTCTGCTCCCGCTGGAGTCCCTGCAATTACGAACCCGACGCTGACGCTGG 356
DB 63 PheAlaYsProProValGlyProLeuAaPheAlaProProGlnProProGlnPro 82
QY 357 GATACCTGGAGAGCAACCTCTACCTTAATTTGCTCCAGAACTCAGAGTGGCTG 416
DB 83 SerGlyValAaPheGlyAaPheProAlaMetCysLeuGlnAaPheAaPheMet 102
QY 417 CTCTTAAGT---CAACAATGCTCAGATGATACCCGAAATTCCGAGTGCAGAAAGC 473
DB 103 AaPheGlnGlyLeuAaPheMetLeuLeuLeuMetProProLeuSerMetSerGlnAaP 122
QY 474 TGGCTTACCTGGAACATCTATGCGCTCCGCGCAGATACAGAGTTCAGAGTCCCCTG 533
DB 123 CysLeuYrLeuAaPheGlnLeuYrThrProThrH1eAaH1eGlnGlyLeuAaPheProVal 142
QY 534 TTGGTGTGTCTCCAGAGAGTGCCTTCAGACCTGGCTCAGGCTCCATCTTGAATGGCTGC 593
DB 143 MetValTrpL1eH1eGlyAlaLeuValAlaGlyMetAlaSerMetYrAaPheGlySer 162
QY 594 GCGCTGGCTGCTATGAGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 653
DB 163 MetLeuAlaH1eGlnAaPheValAlaH1eGlnAaPheGlyLeuAaPheGlyLeu 182
QY 654 GGTTCCTTCCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 713
DB 183 GlyPhePheSerThrGlyAaPheGlnH1eAaAaGlyAaPheGlyLeuAaPheGlyLeu 202
QY 714 GCTGCTGTCTCTGGCTCAGAAAGCATGATGATGATGATGATGATGATGATGATGATGAT 773
DB 203 AlaAlaLeuAaGTrpValGlnGlnAaPheAlaH1eAaPheGlyLeuAaPheGlyLeuVal 222
QY 774 ACCATCTTGGGAGTCCGCGGAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 833
DB 223 ThrLeuPheGlyGlnSerAlaGlyLeuYrThrSerValSerSerH1eAaValAaPheProMet 242
QY 834 GCGAAGGCTTATTCACAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 893
DB 243 SerLysGlyLeuPheH1eGlyAlaH1eMetGlnSerGlyAaPheValMetProGlyLeu 262
QY 894 GAGGCGCATGATTATGAGAAAGTGAAGCATGATGATGATGATGATGATGATGATGATGATG 953
DB 263 IleSerSerSerSerGlnMetVal-----YrThrL1eValAlaAaPheAaPheAaPhe 280
QY 954 AATGGCTCAGACTTGAAGGCTCTGAGAGTCCGAGAGCAAAACCTCCAGAGACTG 1013
DB 281 AlaAlaValAaPheThnGlnThrLeuValSerCysLeuAaGlyLysSerGlnAlaGlnL1e 300
QY 1014 CTGACCTCAGACGCAAAACAAAGCTTTTCACTGAGTGTGTGAATGGTGTCTTCTTCT 1073
DB 301 LeuAlaLeuAaPheValAaPheLysL1eLeuProAlaValAaPheGlyLysLysLysLysLys 320
QY 1074 AATGAGCTCAGACTTATGCTCAGAAAGCATTTAAAGCAATCTCTTCATCATCGGA 1133
DB 1133

```

```

DB 321 LysH1eProGlnGlnLeuLeuAlaSerAlaAaPheAaPheProValProSerL1eLeuGly 340
QY 1134 GTCAATACACAGAGTGTGCTTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1187
DB 341 ValAaPheAaPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 360
QY 1188 CTGAGTGGCTCCAGCAAGTCCCTT-----GCCCTCATTGATGACAAACATCCTCGAC 1241
DB 361 LysGlnL1eThrArgValThrLeuProAlaLeuLysSerThrAlaAaPheGlnMetLys 380
QY 1242 ATCCGCGCTCAGATTTTGAACCTTGTGCTTAATGAATCTTCCATGACAGACATCCCTG 1301
DB 381 LeuProProGlnGlyCysGlyAaPheLeuMetGlnGlyTyrMetGlyAaPheAaPhePro 400
QY 1302 ACTGAAATCCGAGACAGCTTCTGAGCTTGTGGAGATGATGATGATGATGATGATGATGATG 1361
DB 401 GlnThrLeuGlnAlaGlnPheArgGlnMetGlnGlyAaPheThrPheValL1eProSer 420
QY 1362 CTGATCAGACGCTCGATATACAGAGATGATGATGATGATGATGATGATGATGATGATG 1421
DB 421 LeuGlnValAlaH1eAaPheGlnArg---AlaH1eAaPheProValYrPheYrGlnPheGln 439
QY 1422 CACCGGCTCAGTGTGTTGAAGAGAGCGGCTTTGTCAAAGCCGACACGCTGAT 1481
DB 440 H1eAaPheProSerPhePheLysAaPheThrLysPheProH1eValLysAlaAaPheH1eGlyAaP 459
QY 1482 GAGTCCGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1541
DB 460 GlnPhePhePheValPheGlyAaPhe---LeuLeuGlySer-----L1eLysPhe---Ala 475
QY 1542 GCGACGAGAGAGAGAGATTTACAGAGCGGAGAGATGATGAAATACCTGAGCTTGTCT 1601
DB 476 SerThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 495
QY 1602 CGAACCGGAAATCTTAATGGAAGCACTTGTCTGTGAGCGACCTTATATATGATGATGAG 1661
DB 496 ArgH1eGlyAaPheProAaPheSerGlnAaPheProTyrTrpProMetLeuAaPheH1eAaPheGly 515
QY 1662 CAGTACCTCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1721
DB 516 GlnYrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 535
QY 1722 GATTTTGAACAGACACCATCCCC 1745
DB 536 GlnPheThrThrLysL1eLeuPro 543

RESULT 12
Q8BM97 PRELIMINARY; PRT; 559 AA.
AC Q8BM97;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Mus musculus adult male colon cDNA, RIKEN full-length enriched
DE library, clone:9010624L02 product:similar to CARBOXYESTERASE (EC
DE 3.1.1.1) (ALT-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE)
DE (PROCAINE ESTERASE) (METHYLBUTYRASE).
GN Name=9010624L02R1f;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OC NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carinci P, Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;

```



```

OY 810 TCTAGTCTTATCTGCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGGAGAT 869
DB 232 SerSerHisValValSerProMetSerIysGlyLeuPheHisGlyAlaIleMetGlySer 251
OY 870 GGGGTGGCCATCTCTTACCTGAGGCCATGATTAAGAAAGTAGAGAGACCTGACG 929
DB 252 GlyValAlaLeuLeuProTyrLeuIleThrAspThrSerGluMetValSer-----Thr 269
OY 930 GTGGTTGCATTTCTGTGTGAACAAATGGCTGACAGTCTGAGAGCCCTGTCAGAGTCCG 989
DB 270 ThrValAlaIysLeuSerGlyCysGlyAlaMetAspSerGlyAlaLeuValArgCysLeu 289
OY 990 AGGACAAACCTCCAGAGGCTGACCTCAGACCAAGGCTTGCATTCATCGA 1049
DB 290 ArgGlyIysSerGlyAlaGlyIleLeuAlaIleMetIysLeuValGlnMetIleProAla 309
OY 1050 GTGGTTGATGCTCTTCTTCTTATGAGCCTTAGATCTATGTCTGAGAAAGCATTT 1109
DB 310 ValValAspGlyGlyIlePheProArgHisProIysGlyLeuLeuAlaSerGluAspPhe 329
OY 1110 AAAGCAATTCCTTCATCATCGAGATCAATACAGAGATGGCTTCTGCTGCTTACG 1169
DB 330 HisProValProSerIleIleGlyValAlaMetAspGlyIleTyrThrIlePro--- 348
OY 1170 AAGAGGCTCTGAGATCCTCAGTGGCTCCAAAGTCCCTGACCTCATCATGATCAA 1229
DB 349 -----ValValMetGlySerAlaGlnIleIleIysGlyIleThrArgGlu 363
OY 1230 AACATC-----CTGCACATCCCGGCTCAGTATTTG 1259
DB 364 AsnLeuGlnAlaValLeuIysAsnThrThrAlaGlnLeuMetLeuProGluCysSer 383
OY 1260 CACCTTGGGTAATGAATCTTCAATGACAAAGCCTCCTGACATGAATCCGAGCAAT 1319
DB 384 AspLeuLeuMetGluGlyIleIysMetGlyAspThrGlyAspAlaGlnIleThrLeuGln 403
OY 1320 CTTTGGAGCTTCTGAGATGAGTCTTGTGTGCTCCGACCTGATCAAGCTGATAT 1379
DB 404 PheThrGluMetMetGlyIleAspPheMetPheValIleProAlaLeuGlnValAlaTyrPhe 423
OY 1380 CACAGAGATGCTGTGACCTGTCTTCTTATGATTTCCGACCCGCTCAGTCTTT 1439
DB 424 GlnArg-----SerHisAlaSerValTyrPheTyrGluPheGlnHisGlnIleAlaSerLeu 442
OY 1440 GAAGACAGAGACCGGCTTTTGTCAAGCCGACACGCTGATGAGTCCGCTTGTGTTG 1499
DB 443 IysAspValAlaArgProThrHisValIysAlaAspHisAlaAspGlyIleProPheValPhe 462
OY 1500 GGTGTGCTCTC-----CTGAAGGGGACATTGTATGTTGAGAGGACCCGAGGAG 1553
DB 463 GlyTyrPhePheThrAspMetIysLeuAspPhe-----ThrGluGly 476
OY 1554 GAGAGATTACTAGCCGAGAGATGATGAATATCTGGGCTACCTTGTCTGAAACCGGAGAT 1613
DB 477 GlnIysLeuLeuSerHisArgIleMetIleTyrTyrPheAlaAspPheAlaArgHisGlyAsn 496
OY 1614 CCTAATGGAGAACCTGTCTGTGCGCAGCTTATATCTGACATGAGATCTCCAG 1673
DB 497 ProAsnSerGluGlyLeuProTyrTyrProValIleMetAspHisAspGluGlnTyrLeuGln 516
OY 1674 CTGAGCTTGAACATGAGCTCGGACAGACATCAAGAACCGCGGTGATTTTGGAGC 1733
DB 517 LeuAspThrGlnProAlaValGlyArgAlaLeuIysSerArgTyrGluLeuGlnPheTyrPhe 536
OY 1734 AGCACCATCCCC-----CTGATCTGTGCTGCTCCGACATGCTCCACATGCTCTTT 1784
DB 537 IysThrIleuSerGlnIysIleGlnGluIleuArgAlaSerGlnAspArgHisThrGluLeu 556

```

```

DT 05-JUL-2004 (TRENBLREL. 27, last sequence update)
DT 05-JUL-2004 (TRENBLREL. 27, last annotation update)
DE LOC431703 protein (Fragment).
GN Name=LOC431703;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
OX NCBI_TaxId=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Fellinoid R.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodigues R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smallus D.E., Scherch A., Schein J.E.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Klein S., Gernard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/1lipase family.
DB EMBL: BC074230; AAH74230.1;
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR000379; Ser_esterase.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
FT NON_TER 1
SQ SEQUENCE 568 AA; 63032 MW; 9336DDB6B53931A1 CRC64;

```

Alignment Scores:

```

Pred. No.: 2,74e-84 Length: 568
Score: 1263.50 Matches: 258
Percent Similarity: 64.45% Conservative: 87
Best Local Similarity: 48.22% Mismatches: 169
Query Match: 32.63% Indels: 21
DB: 2 Gaps: 8

```

US-10-023-515-1 (1-2158) x Q6GMS4 (1-568)

```

OY 177 GAAAAAAGCTGGGCTTGTGCTGAGAGGCGCAAGAGAACACCAAGCTGGATGATTCAG 236
DB 28 GlyIleTyrGlyThrGluAspAlaArgProLeuLeuThrThrAsnTyrGlyGlnLeuLeu 47

```

```

QY 237 GGCAAGACAGTACATCTGCTGGAAAGCCCTGCTGCTGCAAGCTGTCTCTCGAGCTCC 296
DB 48 GtlystrinValGlyAlaValysglutThrAspArgLeuIleHisValPheMetGlyValPro 67
QY 297 TTGTGCTGCTCCCGCTGGAGATCCCTGCGATTATACGAATCCGACCTCGACATCCGCTCG 356
DB 68 PheAlaValProIleGlyProLeuThrPheGluAspProGlnProProGlnProIlePro 87
QY 357 GATACTTGCCGAGAACCCACCTCCCTCAACCTTAAATTTGTGCTCCAGACATCGAGTGGCTG 416
DB 88 SerSerIleArgGluAlaThrGluAsnProProMetCysLeuGlnAspLysIleGlyMet 107
QY 417 -----CTCTTAGATCAACACATGCTCAAGGAGCATTCACCGGAATTCGAGTGTCAAA 470
DB 108 GluGlnLeuAlaAspPhePheIleAlaValPheAspPheProPro-----ValSerGlu 125
QY 471 GACTGCTCTTACCTGAACATGTATGCGCTGCCACCGCGATACAGGCTCCAACTCCGCC 530
DB 126 AspCysLeuTyrlleuAsnValPheThrProAlaAspArgGlyGluAsnProGlnLeuPro 145
QY 531 GTCTTGAGTGGTCCCGACGAGGTCCTTCAAGATGCTGCTCAAGCTCTCATCTTGTATGG 590
DB 146 ValMetValPheIleHisGlyGlyLeuThrMetGlyAlaGlyMetPheGlnGly 165
QY 591 TCCGCTGCTGCTGCTTATGAGACGTCGTGTTGTGCTGCTCAAGTACCGGCTTGAAGA 650
DB 166 SerAlaLeuSerAlaTyrlleuAsnValValValSerIleGlnTyrlleuGlnGly 185
QY 651 TTGGGTTCTTCCACCAATGGATGAGATGCTCCGGGAACTGGGCTTCAAGACCA 710
DB 186 MetGlyPhePheSerThrGlyAspLysGlyAlaArgGlyAsnTyrlleuPheLeuAspGln 205
QY 711 GTGGCTGCTGTCTGCTGCTCCAGAAACATGCAAGTCTTGGTGGGAGCCCTGCT 770
DB 206 ValAlaAlaLeuArgTrpValArgAspAsnIleLysAspPheGlyLysAsnProGlnSer 225
QY 771 GTACACATCTTTGGCGAGTCCCGCGAGCCATAGTGTCTTACTGTATCTGTCTCC 830
DB 226 ValThrIlePheGlyLeuSerIleGlyGlyLeuSerValSerAlaGlnValLeuSerPro 245
QY 831 ATGGCCAAAGGCTTATTCACAAAGCATCATGAGAGGAGGAGGCTTCAATCCCTTAC 890
DB 246 LeuSerTyrlleuPheHisArgAlaIleSerGlnSerGlyValAlaIleLeuProGly 265
QY 891 CTGGAGCCCAATGATTATGAGAAGTGAAGACCTGCAAGTGTGCAATTTCTGTGGT 950
DB 266 LeuMetAla-----SerLysThrGlyLysIleLeuProIleLeuHisValAla 282
QY 951 AACAAATGCTGAGACTCT---GAGGCGCTGAGAGTGTGCTGAGACAAACCTCCAG 1007
DB 283 AsnIleSerSerCysSerValSerSerLeuAlaAspCysLeuLysLysIleThrGluAsp 302
QY 1008 GAGCTGCTGACCTTCAGCCGAAACAAAG-----TCTTTCACCTGAGTGGTGTGAT 1061
DB 303 GluIleValAlaIleSerAlaIleMetLysPheValAlaPheProAlaValAlaAspGly 322
QY 1062 GCTTCTTCTTCCAAAGAGCTCTATCATCTTGTCTCAAGAAACATTTAAACATTTCT 1121
DB 323 ValPheLeuProLysProAlaGlnGluIleLeuAlaSerLysGluSerAsnProValPro 342
QY 1122 TCCATCATCGAGTCAATACCAAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
DB 343 PheLeuIleGlyValAlaAsnHisGlyIlePheGlyTrpIleLeuProLeuAlaLeuAsn-- 361
QY 1182 GAGATCTCTCAGTGGTCCAAACAAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
DB 362 -----IleSerGlyTyrlleuArgGlnGlyMetGlyLysAspIleGlnSerIleLeuVal 379
QY 1242 ATCCGCTCTCAGTATTTGTCAC-----CTTGGCGCTAATGATAC 1280
DB 380 AlaLeuPro-----PheValHisSerPheThrSerValValProPheIleMetGlnGly 398
QY 1281 TTCATGACAAAGACCTCCCTGAGTGAATCCGAGACAGTCTTCTGAGCTTGTGAGAT 1340

```

```

DB 399 PheGlyAspThrAsnAspProLysGluLeuAspAsnAspPheLeuAspLeuValGlyAsp 418
QY 1341 GTCTTCTTTTGGTCTCCCTGACTGATCACAGCTGCAATATACAGAGATGCTGTGACCT 1400
DB 419 IleIlePheValIleProAlaLeuArgThrAlaLysTyrlleuHisArgAspSerCysIlePro 438
QY 1401 GTCTACTTATGAGTTTGGGACCGGCTCAGTCTGCTTGAAGACACAGACCGGCTTTT 1460
DB 439 ValTyrlleuPheTyrlleuPheGlnPheHisArgProSerMetTyrlleuAspSerLysAspPhe 458
QY 1461 GTCAAGCCGACACAGCTGATGAGTCCGCTTGTGTGCTGCGGAGTCTGCTGCTGCTGCT 1520
DB 459 ValLysAlaAspHisGlyAspGluLeuTyrlleuValGlyGlyProPheLeuLysSer 478
QY 1521 GACATTTTATCTTGAAGAGACCCAGGAGAGAGAAATTACTGAGCCGAGATGATG 1580
DB 479 GlyIleLeuPheLysSerAsnGlyTyrlleuGlnGlnGlyLysIleLeuSerTyrlleuMet 498
QY 1581 AATATGAGGCTACCTTGTGCAACCGGAAATCCTTAATGGGAACGACTGTCTGTGG 1640
DB 499 LysTyrlleuAlaAsnPheAlaArgAsnGlyAspProAsnGlyLeuGlyLeuAlaGlyTrp 518
QY 1641 CCACTTATATCTGACTGACACGACTTCCAGCTGACTGACTGAAATGAGCTCGACAG 1700
DB 519 ProLysTyrlleuAspGluAspGluAspTyrlleuGlnIleLysLeuThrGlnLysSerGln 538
QY 1701 AGACTCAAAAGAACCGGAGTGTGATTTTGGACGACGACCATCCCC 1745
DB 539 ArgLeuLysGlyGlyAlaArgPheLysPheThrValThrLeuPro 553

```

RESULT 15

```

ID 070177 PRELIMINARY; PRT; 561 AA.
AC 070177;
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Cardoxylesterase precursor (EC 3.1.1.1).
GN Name=cardoxylesterase;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Small intestine;
RA Some T., Kunimoto T., Isebe M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the type-B cardoxylesterase/lipase family.
DR EMBL: AB010632; BAA25691.1; -.
DR HSRP: P37967; I0E3.
DR GO: GO:0004091; F:cardoxylesterase activity; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR InterPro: IPR002018; CarboxesteraseB.
DR InterPro: IPR000379; Ser esters.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KM Hydrolyase; Signal.
FT SIGNAL 1
FT CHAIN 26 Potential.
FT CHAIN 27 561 cardoxylesterase.
SQ SEQUENCE 561 AA; 62239 MW; 73A46C33P969398 CRC64;

```

Alignment Scores:

```

Pred. No.: 3,23e-84 Length: 561
Score: 1262.50 Matches: 275
Percent Similarity: 62.41% Conserved: 77
Best Local Similarity: 48.76% Mismatches: 179
Query Match: 32.61% Indels: 33
DB: 2 Gaps: 13

```

US-10-023-515-1 (1-2158) x 070177 (1-561)


```

QY 96 ATGCCAGAGGA---CTTACTTCATCTGCTTCACAAATGGTCTTTTCTGTAATCTCCAG 152
DB 1 MetProhArganGlnMetHisSerTrpLeuMetAlaValLeuPheGlyLeuLeuLeu--- 19
QY 153 CCCCTGTTGGGACACAGACAGTGGGGAAMAACCTGGGCTTCTGTCGAAGGGCCACAGAG 212
DB 20 ---LeuLeuIleHisValGln---GlyGlnAspSerProGlnSer---SerTrpIleArg 36
QY 213 AACACACAGGCTGGGATGGATTCAGGCGAACAGACATCTGCTCTGGGAAGCCCTTGCT 272
DB 37 ThrHisHisThrGlnGlyValGlnGlyLeuAspHisValIleAspThrIleAspGly 56
QY 273 GTGAAGCTGTTCTGGAGATCCCTTTGCTGCTCCCGGCTGGAGATCCCTGAGATTACG 332
DB 57 ValHisThrPheLeuGlyIleProPheAlaIleProValGlyProLeuArgPheAla 76
QY 333 AACCCGACGCTGCATCCGCTGGAGATTAACCTGCGAGAACCCATCCCTAATTTG 392
DB 77 ProProGlnProProGlnProTrpSerGlyValArgAspAlaThrSerGlnProAlaMet 96
QY 393 TGCTTCCGAACATCAGAGTGGCTG-----CTTTAGATCAACATGCTCAAG 440
DB 97 CysLeuGlnAsnLeuAspIleLeuAspGlyValGlyLeuLeuAspMetIleMetIle--- 115
QY 441 GTGCATTACCGGAATTCGAGATGTCAGAGATCGCTCTACTGAAATCTAATCGCT 500
DB 116 -----LeuSerSerIleSerMetSerGlnAspCysLeuTyLeuAsnAlaIlePro 133
QY 501 GCCACGCGGATACAGAGCTCCAAAGCTCCCGTCTTGATGATGTTCCACAGAGTCCCTTC 560
DB 134 AlaHisAlaArgGlnGlySerHisLeuProValMetValTrpIleHisGlyIleAlaLeu 153
QY 561 AAGATGCTCAGAGCTCCATCTTTATATGAGGTCGCGCTGGCTGCTATGAGAGCTGCTG 620
DB 154 ValValGlyMetAlaSerMetCysArgGlySerLeuLeuThrValAsnGlyAspLeuVal 173
QY 621 GTTGTGTCGTCGACGATCCGCTAGGAATATTTGGTTTCTTCAACACATGGATACAGAT 680
DB 174 ValValThrIleGlnTyArgLeuGlyValLeuGlyPhePheSerThrGlyAspGlnHis 193
QY 681 GCTCCGAGGAACAGCTGGCTTCAAGACAGACAGAGTGGCTGCTGCTGGGTCAGAGAAC 740
DB 194 AlaArgGlyAsnTrpGlyTyLeuAspGlnValAlaAlaLeuAspGlyTrpValGlnGlnHis 213
QY 741 ATCGAGTCTTGGTGGGAGACCCCAAGCTCTGTGACATCTTTGGGAGTCCGCGGAGCC 800
DB 214 IleAlaHisPheGlyGlyAsnProAsnArgValThrIlePheGlyGlnSerAlaGlyIle 233
QY 801 ATAAATGTTTCTAGCTTATATCTGTCCCAATGCGCAAGGCTTATCCACAAAGCCATC 860
DB 234 ThrSerValSerSerHisValIleSerProMetSerGlnGlyLeuPheHisGlyAlaIle 253
QY 861 ATGAGAGTGGGAGGAGCCATCCCTTAACCTGAGAGCCCATGATTAAGAAAGATGAG 920
DB 254 MetGlnSerGlyValAlaLeuLeuProAspLeuIleSer-----GlnHisSerGln 270
QY 921 GACSTGCAG---GTGATTCACATTTCTGTGTAACAAATGCGTCAAGCTGAGGCGCTG 977
DB 271 ThrValSerThrThrValAlaIleAspSerGlyCysGlnAlaMetAspSerGlnAlaLeu 290
QY 978 CTGAGAGTGCCTGAGACAAACCCCTCCAAAGAGCTGAGCTGACCTCAAGCAAGAAACAG 1037
DB 291 ValArgCysLeuAlaGlyAlaLeuSerGlyAlaGlnIleLeuValIleAsnIleValPheLeu 310
QY 1038 TCTTCACTGAGAGTGGATGATGCTTCTTCTTAATGAGCTCTAGACTATATGCT 1097
DB 311 MetIleProAlaValAlaAspGlyGlyPheLeuProAlaGlnIleProGlyGlyLeuAla 330
QY 1098 CAGAAAGCATTTAAAGCAATTTCTTTCATGATCGAGATGCAATACCAAGATGAGGCTTC 1157
DB 331 SerGlnAspPheHisProValProSerIleIleGlyValAlaMetThrAspGlyIleCysCys 350

```

```

QY 1158 CTGCTGCTCATG-----AAGAGAGCTCTGAGATCTCTAGT 1193
DB 351 ThrIleProMetValMetGlyTyThrAlaGlnIleIleLeuGln-----LeuSer 366
QY 1194 GAGTCCAAACAGTCCCTTGGCTCCATGATTAACAAACATCTGACATCCCGCTCAG 1253
DB 367 ArgGlnAsnLeuGlnAlaValLeuIleAspThrAlaAlaGlnMetMetLeuProGln 386
QY 1254 TATTTGACATCTTGTGGCTAATGATATCTTCAATGACAGACATCCCTGATGAATCCGA 1313
DB 387 CysGlyAspLeuLeuMetGlnGlyIleIleCysGlnAsnThrAspAspSerGlnThrLeuGln 406
QY 1314 GACAGCTCTTGCAGCTTCTGAGATGATGTTCTTTGTGTCCTCCGACATGATCAAGCT 1373
DB 407 IleGlnTyTyThrGlnMetMetCysAspPheLeuPheValIleProAlaLeuGlnValAla 426
QY 1374 CGATATCAACAGATGCTGTGACCTGTCTATCTTATGAGTTTGGACCGGCTCAG 1433
DB 427 HisPheGlnArg---SerHisAlaProValTyPheTyGlnPheGlnHisIleAlaProSer 445
QY 1434 TGCTTTGAAGACAGAACCCGCTTTGTCAAGCCGACCAACCGTATGAAGTCCGCTT 1493
DB 446 TyPheLeuAsnValArgProHisValIleValAlaAspHisValAlaAspGlyValProPhe 465
QY 1494 GTGTCGGTGGTGCCTTCTCTGAAGGGGAGCATGTTATGTTTGAAGAGCCACGAGAG 1553
DB 466 ValPhe---GlySerPhePheSerGlyMetCysLeuAspPhe-----ThrIleGln 481
QY 1554 GAGAAGTTACTGAGCCGGAAGATGTAATATCTGGGCTTACTTGTGGAACCGGAGAT 1613
DB 482 GlnArgLeuLeuSerArgArgMetMetCysTyTyTrpAlaAsnPheAlaArgGlnGlnAsn 501
QY 1614 CCTAATGGGAAGACCTGTCTCTGTGGCCACTTATATGACATGACAGACATACCTCCAG 1673
DB 502 ProAsnSerGlnGlyLeuProTyTyTrpProAlaLeuAspHisAspGlnGlnTyLeuGln 521
QY 1674 CTGACCTTGAACATGAGGCTCGGACAGAGACTCAAGAGACCGGCGGATGATTTGAGAC 1733
DB 522 LeuAspThrHisProAlaValAlaAspArgAlaLeuIleValAlaArgLeuGlnPheThrHis 541
QY 1734 AGCACCATCCCC 1745
DB 542 LysThrLeuPro 545

```

Search completed: June 15, 2005, 08:52:45
 Job time : 226.639 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 07:48:14 ; Search time 32.6132 Seconds

(without alignments)
9878.981 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 3872

Sequence: 1 ccaagcgcgcgaacacagc.....aaaaaaaaaaaaaaaaaa 2158

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model DEV=xlh
-O=/cgn2_1/USPTO_gpool/6664091/runat_14062005_133500_9360/app_query.fasta_1.4238
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODER=LOCAL -OUTFMT=pco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091 @CGN 1.1 54 @runat_14062005_133500_9360 -NCP=6 -ICU=3 -NO_MMAB
-LARGEQUERY -NEG_SCORES=0 -WAIT -DESBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3079	79.5	581	4	US-10-023-515-2
2	1230	31.8	559	4	US-09-595-682B-28
3	1230	31.8	559	4	US-09-949-016-6426
4	1230	31.8	577	4	US-09-949-016-9670
5	1150.5	29.7	565	4	US-09-595-682B-21
6	1138.5	29.4	566	3	US-09-264-737-2
7	1130.5	29.2	543	4	US-09-595-682B-26
8	1091.5	28.2	584	2	US-08-845-295A-2
9	1091.5	28.2	584	3	US-09-140-933-2
10	1091.5	28.2	584	3	US-09-146-661-2
11	1091.5	28.2	584	3	US-09-150-515-2
12	1034.5	26.7	539	3	US-09-264-737-1

13	942.5	24.3	454	3	US-08-446-100-26	Sequence 26, Appl
14	942.5	24.3	454	3	US-08-446-100-28	Sequence 28, Appl
15	942.5	24.3	454	3	US-08-446-100-30	Sequence 30, Appl
16	942.5	24.3	454	3	US-08-446-100-31	Sequence 31, Appl
17	940.5	24.3	454	3	US-08-446-100-27	Sequence 27, Appl
18	940.5	24.3	454	3	US-08-446-100-29	Sequence 29, Appl
19	871	22.5	574	4	US-10-023-515-4	Sequence 4, Appl1
20	800.5	20.7	933	4	US-09-949-016-8386	Sequence 8386, Ap
21	789.5	20.4	933	4	US-09-949-016-8387	Sequence 8387, Ap
22	786	20.3	823	4	US-09-949-016-6888	Sequence 6888, Ap
23	777.5	20.1	843	4	US-09-491-356C-20	Sequence 20, Appl
24	777	20.1	575	1	US-08-348-920-1	Sequence 1, Appl1
25	775.5	20.0	823	4	US-09-491-356C-23	Sequence 23, Appl1
26	774	20.0	575	1	US-08-348-920-2	Sequence 2, Appl1
27	772	19.9	848	4	US-09-491-356C-22	Sequence 22, Appl
28	769	19.9	836	4	US-09-491-356C-21	Sequence 21, Appl
29	765.5	19.8	614	3	US-08-446-100-25	Sequence 25, Appl
30	764.5	19.7	614	1	US-07-732-962A-2	Sequence 2, Appl1
31	764.5	19.7	614	2	US-08-370-156-2	Sequence 2, Appl1
32	764.5	19.7	614	3	US-08-446-100-19	Sequence 19, Appl
33	764.5	19.7	614	3	US-08-446-100-21	Sequence 21, Appl
34	764.5	19.7	614	3	US-08-814-095-2	Sequence 2, Appl1
35	764.5	19.7	614	5	PCT-US92-06106-2	Sequence 2, Appl1
36	764.5	19.7	645	4	US-09-949-016-7063	Sequence 7063, Ap
37	764.5	19.7	645	4	US-09-949-016-7064	Sequence 7064, Ap
38	762.5	19.7	614	3	US-08-446-100-20	Sequence 20, Appl
39	761.5	19.7	614	3	US-08-446-100-23	Sequence 23, Appl
40	760.5	19.6	614	3	US-08-446-100-22	Sequence 22, Appl
41	759.5	19.6	600	2	US-08-370-156-4	Sequence 4, Appl1
42	759.5	19.6	600	3	US-08-814-095-4	Sequence 4, Appl1
43	759.5	19.6	600	3	US-08-975-084-1	Sequence 1, Appl1
44	755.5	19.5	617	2	US-08-370-156-6	Sequence 6, Appl1
45	755.5	19.5	617	3	US-08-814-095-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curiis, Rory A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2

Alignment Scores:
Pred. No.: 2.11e-292
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 79.52%
DB: 4
Caps: 0
US-10-023-515-1 (1-2158) x US-10-023-515-2 (1-581)
QY 96 ATGCACAGGAGTACTTCACTGCTTCACAAATGTCCTTTCTGATTCACAGCC 155
|||||

```
Dh 1 MetProGInglyLeuThrSerSerAlaSerGIntrProCy8PhePheLeuIleLeuGInPro 20
Qy 156 CTGTGGGACACAGACAGTGGGGAAAACTGGGCTTTGTCTGAAGGGCCACAGAGAAC 215
Dh 21 LeuLeuGInlySArGIntrProGlySerThrGlyProSerAlaGInGlyProGInArGAsn 40
Qy 216 ACCAGGCTGGATGATTCAAGGCAAGCAGATGCTGGGGAGCCCTGGCTGTG 275
Dh 41 ThrArgLeuGInlyTrpIleGInGlySerGInValThrValLeuGInlySerProValProVal 60
Qy 276 AACGTTCTCTCGAGAGTCCCTTTGTCTGCTCCCGCTGGAGTCCCTGCATTTACGAC 335
Dh 61 AenValPheLeuGInlyValProPheAlaIaIaProProLeuGInlySerLeuArGpHeThrAsn 80
Qy 336 CCGAGCCTGCATGCGCCCTGGAGATACCTTGGAGAACCCACTCTACCTTAATTTGTGC 395
Dh 81 ProGInProAlaSerProTrpAspAsnLeuArgGInIaThrSerIyProAsnLeuCy8 100
Qy 396 CTCCAGAACTCAGAGTGGCTCTTATGATCAACACATGCTCAAGGTGATTAACCCGAAA 455
Dh 101 LeuGInAsnSerGIntrPLeuLeuLeuAspGInhiSerLeuIySerValhiSerIyProIy8 120
Qy 456 TTGGAGATGTCAGAGACCTGCTCTTACCTGAACATCTATGCGCTGCCACGCCATPACA 515
Dh 121 PheGInlyAlaSerGInlyAspCy8LeuIyTrLeuAsnIleTyAlaProAlaIaIaAspThr 140
Qy 516 GGCTCCAGAGTCCCGCTTTGTGTGTGTCTCCAGAGGTGCTTCAAGATGGCTCAGGC 575
Dh 141 GlySerIySerLeuProValIleuValTrpPheProGlyGlyAlaPheIySerThrGlySerAla 160
Qy 576 TCATCTTTGATGGTCCCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 635
Dh 161 SerIlePheAspGInlySerAlaLeuAlaIaTyGInAspValIleuValIaIaGIn 180
Qy 636 TACCGCTAGGAATATTTGGTTTCTTCAACACATGGAGATCAGATCCTCCGGGAACTGC 695
Dh 181 TyArgLeuGInlyIlePheGInlyPhePheThrTrpAspGInhiIaIaProGInlyAsnTrp 200
Qy 696 GCCTTAAAGACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
Dh 201 AlaPheIyAspGInValIaIaIaLeuSerTrpValGInIyAsnIleGInPhePheGInly 220
Qy 756 GGGGACCCGAGCTGTGACCATCTTTGGCGAGTCCGGGGACCATAGTGTTCATAGT 815
Dh 221 GlyAspProSerSerValThrIlePheGInlyIaIaIaIaIaIaIaIaIaIaIaIaIa 240
Qy 816 CTTTAACTGTCTCCATGCGCAAGAGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 875
Dh 241 LeuIleLeuSerProMetAlaIySerGInlyLeuPheIaIaIaIaIaIaIaIaIaIaIaIa 260
Qy 876 GCCATCATCCCTTACCTGAGGCGCCATATTTAGAGAGAGTGAAGCTGCTGAGTGGT 935
Dh 261 AlaIleIleProIyTrLeuGInIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 280
Qy 936 GCACATTTCTGTGTGAACATGCGCTCAGACTCTGAGGCGCTGAGAGTGGCTGAGACA 995
Dh 281 AlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 300
Qy 996 AAACCCCTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
Dh 301 LySerProSerIyValLeuLeuThrLeuSerGInlySerThrIySerPheThrArgValIa 320
Qy 1056 GATGATGCTTTCTTTCTTATGAGCTCTGAGTCTATGCTGCTGCTGCTGCTGCTGCTG 1115
Dh 321 AspGInlyAlaPhePheProAsnGInProLeuAspLeuLeuSerGInlyValaPheIyVala 340
Qy 1116 ATTGCTTCCATCATGAGAGTCAATTAACCAAGAGTGGCTTCTGCTGCTGCTGCTGCTG 1175
Dh 341 IleProSerIleIleGInlyValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 360
Qy 1176 GCTCTGAGATCTCAGTGGCTCAACAGATCCCTTGCCTCATCTGATCAAAACATC 1235
Dh 361 AlaProIuIleLeuSerGInlySerAsnIySerLeuAlaLeuIaIaIaIaIaIaIaIaIa 380

Qy 1236 CTGACATCCCGCTCAGATTTTGCACCTTTGGCTTAATGAATCTTCCATGACAGAC 1295
Dh 381 LeuIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 400
Qy 1296 TCCCTGATGAATCCGAGACAGTCTTGTGACTTGTGAGATGATGATGATGATGATG 1355
Dh 401 SerLeuThrGInIleArGAspSerIleuAspLeuGInlyAspValIaPhePheValIa 420
Qy 1356 CCGTGAATGATACAGCTCGATATACAGAGATGCTGTGACCTGTCTACTTATGAG 1415
Dh 421 ProAlaLeuIleThrAlaArgTyTrhiSArGAspAlaGInIaIaIaIaIaIaIaIaIaIa 440
Qy 1416 TTTCGGACCCGCTCAGTGTGTTGAAGACAGAACCGGCTTTTGTCAAGCCGACAC 1475
Dh 441 PheArgGInSArGProGInCy8PheGInlyAspThrIySerProAlaPheValIyValaAsp 460
Qy 1476 GCTGATGAATCCCGCTTTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1535
Dh 461 AlaAspGInValaArgPheValaPheGInGlyAlaPheLeuIySerGInlyAspIleValaMetPhe 480
Qy 1536 GAAGAGCCACAGAGAGAGAGATTACTGACCCGGAGATGATGAATCTGGGCTACC 1595
Dh 481 GInGInlyAlaThrGInGInGInGInlyIyLeuLeuSerArgIySerMetIySerTyTrpAlaThr 500
Qy 1596 TTGTGCTGAACCGGGAATCCTTAATGGAGACACCTGTCTGTGGCCAGCTTAATCTG 1655
Dh 501 PheAlaArgThrGInyAsnProAsnGInyAsnAspLeuSerLeuIyProAlaTyRAsnLeu 520
Qy 1656 ACTGAGACATCCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1715
Dh 521 ThrGInGInlyTrLeuGInIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 540
Qy 1716 CGGCTGATTTTGAACACAGACCATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTG 1775
Dh 541 ArgValaAspPheThrPheThrIleProLeuIleLeuSerAlaSerAspMetLeuIaIa 560
Qy 1776 AGCCCTCTTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
Dh 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGInProPhePhePheCy8Ala 580
Qy 1836 CCT 1838
Dh 581 Pro 581

RESULT 2
US-09-595-682B-28
/ Sequence 28, Application US/09595682B
/ Patent No. 6800483
/ GENERAL INFORMATION:
/ APPLICANT: Danks, Mary K.
/ APPLICANT: Potter, Philip M.
/ APPLICANT: Houghton, Peter J.
/ TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
/ FILE REFERENCE: SJ-0005
/ CURRENT APPLICATION NUMBER: US/09/595,682B
/ PRIOR FILING DATE: 2000-01-16
/ PRIOR APPLICATION NUMBER: 60/075,258
/ PRIOR FILING DATE: 1998-02-19
/ PRIOR FILING DATE: 1999-02-12
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 28
/ LENGTH: 559
/ TYPE: prt
/ ORGANISM: Homo sapiens
US-09-595-682B-28

Alignment Scores: 2,7e-111 Length: 559
Pred. No.: 1230.00 Matches: 248
```

Percent Similarity: 63.84% Conservative: 91
 Best Local Similarity: 46.70% Mismatches: 168
 Query Match: 31.77% Indels: 24
 DB: 4 Gaps: 8

US-10-023-515-1 (1-2158) x US-09-595-662B-28 (1-559)

```

Oy 186 GGGCTTGTGTAAGGCGCAGAGAAACACAGAGCTGGATTCAGGGCAAGAA 245
Db 26 GAGTGAATGTAAGGCGCAGAGAAACACAGAGCTGGATTCAGGGCAAGAA 45
Oy 246 GTCACTGTGTAAGGCGCAGAGAAACACAGAGCTGGATTCAGGGCAAGAA 305
Db 46 ValHsValysGlyAlaAsnAlaGlyAlaGlnThrPheLeuGlyIleProPheAlaLys 65
Oy 306 CCCCCGTGGAGATCCCTGCGATTTACGAACCCGAGCCGTGATCCCTGGATTAATTTG 365
Db 66 ProProLeuGlyProLeuAlaPheAlaProProGluProGluSerTrpSerGlyAla 85
Oy 366 CGAAGAGCCAGCTCTACCTTAATTTGTGCTCCAGAAC-----TCAGAG 410
Db 86 ArgAsnGlyThrThrHisProAlaMetCysLeuGlnAspLeuThrAlaValGluSerGlu 105
Oy 411 TGGCTGCTTAAGATCAACACATGCTCAAGGTGATTCACCGAAATTCGAGGTGACAGA 470
Db 106 PheLeu-----SerGlnPheAsnMetThrPheProSerAspSerMetSerGlu 121
Oy 471 GACTGCTCTACCTTAAGATCAACATGCTCAAGGTGATTCACCGAAATTCGAGGTGACAGA 530
Db 122 AspCysLeuGlyLeuSerIleThrProAlaHisSerHisGlyGlnSerAsnLeuPro 141
Oy 531 GTCTGTGATGAGTCCAGAGAGTGCCTTCAAGAGCTGATCCCTCAATCTTGTATGGG 590
Db 142 ValMetValIlePheIleGlyGlyAlaLeuValPheGlyMetAlaSerLeuGlyTrpAspGly 161
Oy 591 TCCGCGCTGCTGCTATGAGAGACGTGCTGTTGTGTCGTGTCAGTACCGGCTAGAAAT 650
Db 162 SerMetLeuAlaAlaLeuGlnAsnValValValIleIleGlnThrArgLeuGlyAla 181
Oy 651 TTTGGTTCTTCAACACATGAGATGATGCTCCGGGAGATCGGCTTCAAGAGACAG 710
Db 182 LeuGlyPhePheSerThrGlyAspLysHisAlaThrGlyAsnTrpGlyTyrLeuAspGln 201
Oy 711 GTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
Db 202 ValAlaAlaAlaLeuAspGlyTrpValGlnGlnAsnIleAlaHisPheGlyGlyAsnProAspArg 221
Oy 771 GTGACATCTTTGCGAGATCCGCGGAGACCATAGTGTTCCTTACGTTATCTGCTCC 830
Db 222 ValThrIlePheGlyLeuSerHisGlyIleThrSerValSerSerLeuValValSerPro 241
Oy 831 ATGGCGAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGGCGCATCCCTTAC 890
Db 242 IleserGlnGlyLeuPheHisGlyAlaIleMetGlnSerGlyValAlaLeuLeuProGly 261
Oy 891 CTGAGAGCGCCATGATTAAGAGAGTGAAGCTG-----CAGGTGTTGCAATTTTC 944
Db 262 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaAsnLeu 277
Oy 945 TGTGTAACAATGCTGAGACTGTGAGGCTGCTGAGGCTGCTGAGACAAACCTTCC 1004
Db 278 SerAlaCysAspGlnValAspSerGlnAlaLeuValGlyCysLeuArgGlyLysSerLys 297
Oy 1005 AAGAGCTGTCGACCTTCAACCGCAAGAAACAAAGTCTTTCATCGAGTGTGATGCTGCT 1064
Db 298 GluGlnIleLeuAlaIleAsnLysProPheLysMetIleProGlyValValAlaAspGlyAla 317
Oy 1065 TTCTTTCTAATGAGCTTGAATCTATTTGTCTAGAAAGCATTAAGCAATTCCTTCC 1124
Db 318 PheLeuProAlaGlnHisProGlnIleLeuLeuAlaSerAlaAspPheGlnProValProSer 337
Oy 1125 ATCATCGAGTCAATATACACAGAGTGTGCTTCTGCTGCTC-----ATGAAG 1172

```

```

Db 338 IleValGlyValAsnAlaAsnGluPheGlyTrpLeuIleProLysValMetArgIleThr 357
Oy 1173 GAGGCTCTGAGATCCATGATGCTGCTCAACAAAGTCCCTGCTCATCTGATCAAAAC 1232
Db 358 AspThrGlnLysGlnMetAspArgGluAlaSerGlnAlaAlaLeuGlnLysMetLeuThr 377
Oy 1233 ATCTGCAATCCCGCTCAGATTTTGAACCTTTGGTGAATGAATCTTCATCAAGCAAG 1292
Db 378 LeuLeuMetLeuProProThrPheGlyAspLeuAlaArgGlnGlyTrpIleGlyAspAsn 397
Oy 1293 CACTCCCTGACTGAATATCCAGACAGTCTTTCGACTTTCGAGATGTGTTCTTGTG 1352
Db 398 GlyAspProGlnThrLeuGlnAlaGlnPheGlnGlnMetMetAlaAspSerMetPheVal 417
Oy 1353 GTCCCTGACTGATACAGCTCGATGATACAGAGATGCGGTGACCTGTCTACTTCTAT 1412
Db 418 IleProAlaLeuGlnValAlaHisPhe--GlnCysSerArgAlaProAlaTrpThrGly 436
Oy 1413 GAGTTTCGACACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTTGTCAACCGGAC 1472
Db 437 GluPheGlnHisGlnProSerTrpLeuLysAsnIleArgProHisMetLysAlaAsp 456
Oy 1473 CAGCTGATGAATCCGCTTGTGTTCGAGTGTGCTTCTGAGGGGACATTTGATG 1532
Db 457 HisGlyAspGlnLeuProPheValPhe--ArgSerPhePheGlyGlyAsnTrpIleLys 475
Oy 1533 TTGGAAGAGCGCACGAGAGAGAGAGATTAATCAAGCCGAGAGATGAATGAATCGGGCT 1592
Db 476 Phe-----ThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 492
Oy 1593 ACCTTGTCTGCAACCGGAGATCTTAATGGAACGACGTGTCTGTGCGACGACTTAAT 1652
Db 493 AsnPheAlaArgAsnGlyAsnProAsnGlyGlnGlyLeuProHisIleTrpProLeuPheAsp 512
Oy 1653 CTGACTGAGACGATCTTCAGCTGACTTGAACATGAGCTTGCAGACAGACTCAAGAA 1712
Db 513 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 532
Oy 1713 CCGGCGGTGAGATTTTGAACGACGACCATCCCC 1745
Db 533 HisArgLeuGlnPheTrpLysAlaLeuPro 543

RESULT 3
US-09-949-016-6426
; Sequence 6426, Application US/09949016
; Patent No. 681235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH TYPE 2 DIABETES MELLITUS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 6426
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6426

Alignment Scores:
Pred. No.: 2,76-111 Length: 559
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Mismatches: 91
Best Local Similarity: 46.70% Indels: 168
Query Match: 31.77%

```

```

DB:      4      Gaps:      8
US-10-023-515-1 (1-2158) x US-09-949-016-6426 (1-559)
OY      186 GGGGCTTGTCTGTAAGGGCCACAGAGAAACACAGAGTGGATGATTCAGAGGCAAGCA 245
DB      26 GtGtInAmpSerAlaSerProIleArgThrThrIleThrIleGlnValLeuGlySerIleu 45
OY      246 GTCACTGTGCTGGAGCCCTGTGCTGTAAGACGTGTTCTCTGAGAGTCCCTTGTGCT 305
DB      46 ValHISValIySgLIAlaAmnIaGtIyAlGlnThrPheLeuGlyIleProPheAlaIyS 65
OY      306 CCCCCCTGGAGATCCCTGCGATTATGAAACCCGAGCCGTCGATGAGCCCTGGAGATACGTG 365
DB      66 ProPheLeuGlyProLeuAlaArgPheIleAArgProProGluProGluSerIleProGlyVal 85
OY      366 CGAAGAGCCACCTCTACCTTAATTGTGCTCCAGAAC-----TCAGAG 410
DB      86 ArgAmpGlyThrThrIleAlaSerProIleAmetCylLeuGlnAmpLeuThrAlaValGluSerGlu 105
OY      411 TGAGTGTCTTGAATCAACATGCTCAAGGTGATGACCGAATTCGAGAGTCCAGAA 470
DB      106 PheLeu-----SerGlnPheAmnMetThrPheProSerAmpSerMetSerGln 121
OY      471 GAGTCCCTTACTGAAACATGATATGAGCCCTGCGCCACGATACAGAGCTCCAAAGTCCCC 530
DB      122 AspCylLeuIyIleuSerIleIyThrProAlaHISerHISgLIuGlySerAmpLeuPro 141
OY      531 GTCTTGAGTGTGTTCCAGAGAGGTGCTTCAAGACTGGCCACGTCGATCTTGAATGG 590
DB      142 ValMetValIleIleHISgLIyAlaLeuValPheGlyMetAlaSerLeuIyIleAmpGly 161
OY      591 TCGGCGCTGCTGCTTATGAAACGTGCTGTTGTGCTGCTCCAGTACCGGCTAGAGATA 650
DB      162 SerMetLeuAlaIleuGlnAmpValIleValIleIleGlnIyIleAmpGlyVal 181
OY      651 TTGAGTGTCTTCAACATGAGATGAGAGTCCGAGGAGACGTCGCTTCAAGAGCAG 710
DB      182 LeuGlyPhePheSerThrIyAmpIySgLIyAlaThrGlyAmpIleProIyIleuAmpGln 201
OY      711 GTGGCTGCTGTCTGCTGGGCTCGAAGAACATGAGATTTCTGAGGAGCCCGACCTCT 770
DB      202 ValAlaAlaLeuAArgIyIleValGlnGlnAmpIleAlaHISgLIyAlaAmpProAArg 221
OY      771 GTGACCATCATTTGGGAGTCCGGGAGGACATGATGTTTCACTTAACTGTGCTCC 830
DB      222 ValThrIlePheGlyIleuSerAlaGlyIyIleIleSerValSerSerIleuValIleSerPro 241
OY      831 ATGGCCAAAGGCTTATTCACAAAGCATGAGAGTGGGAGGAGCCATACCTTAC 890
DB      242 IISerGlnGlyLeuPheHISgLIyAlaIleMetGluSerGlyValAlaLeuLeuProGly 261
OY      891 CTGAGAGGCCCATGATTAGAGAAAGTGAAGACCTG-----CAGGTGTTCACATTTTC 944
DB      262 LeuIleAla-----SerSerIleAmpValIleSerThrValIleAlaAlaLeu 277
OY      945 TGTGTGTAACAATGCGCTCAGACTCTGAGGCTGCTGAGGCTGCTGAGAGCAAAACCTTCC 1004
DB      278 SerAlaCylAmpGlnValAmpSerGlnAlaLeuValGlyCylSerAmpGlyIySerIyS 297
OY      1005 AAGAGCTGCTGACCTTACGCCCAAGAAACAAAGCTTTTCACTGAGTGTGATGATGCT 1064
DB      298 GtGtGtIleLeuAlaIleAmpIySgLIyAmpPheIleProGlyValIleAmpGlyVal 317
OY      1065 TTCTTCTTAATGAGCTCTGATCTATGCTCAGAAAGATTTAAAGCATTTCTTCC 1124
DB      318 PheLeuProAArgHISProGlnGlnIleuLeuAlaSerIleAmpPheGlnProAArg 337
OY      1125 ATCATCGAGTCAATTAACACGAGTGTGCTTCTGCTGCT-----ATGAG 1172
DB      338 IISgLIyAlaIleAmpAmpGlnPheGlyIyIleProIleProIySValMetAArgIleIy 357
OY      1173 GAGGCTCTGAGATCTCTCAGTGGCTCAACAGACGTCCTTCCATCTGATACAAAC 1232

```

```

DB      358 AspThrGlnIySgLIyMetAArgGlyAlaSerGlnAlaAlaLeuGlnIyMetLeuThr 377
OY      1233 ATCTTCACATATCCCGCTCAGTATTTGACCTTGTGGTAAATGAATCTTCATGACAAAG 1292
DB      378 LeuLeuMetLeuProProThrPheGlyAmpLeuAlaGlyGlnIyIleGlyAmpAArg 397
OY      1293 CACTCCCTCAGTGAATCCGAGACAGACAGTCTTTCGACTTGTGAGATGTGTTCTTGTG 1352
DB      398 GlyAmpProGlnThrLeuGlnAlaGlnPheGlnGlnMetAlaAmpSerMetPheVal 417
OY      1353 GTCCCTGCACTGATCAGCTCGATATCAGAGATGCTGTGCAACCTGCTTACTTCTAT 1412
DB      418 IISgLIyAlaLeuGlnValAlaHISgLIyPhe--GlnCylSerAlaArgAlaProValIyPheIy 436
OY      1413 GAGTTTGGGACCCGCGCTCAGTCTTGAACACAGAACCCGCTTTTGTCAAGCCGAC 1472
DB      437 GtPheGlnHISgLIyPheProSerIyIleuIySgLIyAmpIleArgProProHISmetIyAlaAmp 456
OY      1473 CAGCGTGAATGAAGTCCGCTTGTGCTGCTGCTGCTTCCGAAAGGCGCATTTGATG 1532
DB      457 HISgLIyAmpGlyIleuProPheValPhe--ArgSerPhePheGlyIleAmpIyIleIyS 475
OY      1533 TTGAAAGAGCCACGAGAGAGAGAAAGTACTGAGCCGGAAGATGATGAATACTGGGCT 1592
DB      476 Phe-----ThrGlnGlnGlnGlnGlnLeuSerArgIySgLIyMetCylIyIleProAla 492
OY      1593 ACCTTGTGCTGAAACCGGAAATCTTAATGGAAGACGCTGTCTGTGCGCAAGTAAAT 1652
DB      493 AampPheAlaArgAmpGlyAmpProHISgLIyGlnIyLeuProHISgLIyIlePheAmp 512
OY      1653 CTGACTGACGATGATCTTCCAGCTGAGTGAACATGAGCCTCGACAGAGATCTCAAGAA 1712
DB      513 GlnIleuGlnIyIleuGlnIleuAmpLeuGlnProAlaValGlyIyArgAlaLeuIySgLIy 532
OY      1713 CCGGAGGTGATTTTGGACACAGACATCCCC 1745
DB      533 HISgLIyLeuIlePheIyIlePheIySgLIyAlaLeuPro 543

```

```

RESULT 4
US-09-949-016-9670
? Sequence 9670, Application US/09949016
? Patent No. 6812339
? GENERAL INFORMATION:
? APPLICANT: VENTNER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 9670
? LENGTH: 577
? TYPE: PRT
? ORGANISM: Human
US-09-949-016-9670

Alignment Scores:
Pred. No.: 2.76e-111
Score: 1230.00
Percent Similarity: 63.84%
Match: 248
Beet Local Similarity: 46.70%
Mismatch: 91
Query Match: 31.77%
Indel: 168
DB: 4
Gaps: 8
US-10-023-515-1 (1-2158) x US-09-949-016-9670 (1-577)

```



```

Oy 186 GGGCTTGTGTAAGGGCAACAGAGAAACACAGAGTGGATTCAGGGCAAGCA 245
Db 44 GlylnaspSerAlaSerProIleArgThrThrIleThrGlnValLeuGlySerIleu 63
Oy 246 GTACATGTGTGGAGAGCCCTGTGCTGTGAACGTGTTCTTGAGATCCCTTGTGCT 305
Db 64 ValHisValValGlyAlaHisAlaGlyValGlnThrPheLeuGlyIleProPheAlaIys 83
Oy 306 CCCCCGTGGAGATCCCTGTGCTGTGAACCGGAGCCCTGTGATCCCTGGATTAATG 365
Db 84 ProProLeuGlyProLeuArgPheAlaProProGluProGluSerIleProGlyVal 103
Oy 366 CGAGAGAGCCACTCTACCTTAATTTGTGCTCCAGAGC-----TCAGAG 410
Db 104 ArgAspGlyThrThrIleProAlaMetCysLeuGlnAspLeuThrAlaValGlySerGln 123
Oy 411 TGGCTGTCTTATAGTCAACATGTCTCAAGGTGATTAACCCGAAATTCGAGTGTCA 470
Db 124 PheIleu-----SerGlnPheAsnMetThrPheProSerAspSerMetSerGln 139
Oy 471 GATGCTCTTACCTTAACATCTATGCGCTGCGCCACCGGATACAGCTCCAGCTCCC 530
Db 140 AspCysLeuGlyLeuSerIleGlyThrProAlaHisSerHisGlySerAsnLeuPro 159
Oy 531 GCTTGTGTGGTGTCCAGAGAGTGGCTTCAAGAGCTGCTGCTCCATCTTATGAGG 590
Db 160 ValMetValTrpIleHisGlyValAlaLeuValPheGlyMetAlaSerLeuGlyArgGly 179
Oy 591 TCCGCGCTGTGCTTATAGAGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 650
Db 180 SerMetLeuAlaAlaLeuGlnAlaValAlaValIleIleGlnThrArgLeuGlyVal 199
Oy 651 TTGGTGTCTTCAACATGGAGTCAAGAGTGTCCGGGAGTGGGCTTCAAGAGCA 710
Db 200 LeuGlyPhePheSerThrGlyAspIleHisAlaThrGlyAsnTrpGlyTyrLeuAspGln 219
Oy 711 GTGGCTGTGTCTGTCTGTCCAGAGAAACATCGAGTCTTCCGTGGGAGACCCGAGCT 770
Db 220 ValAlaAlaLeuArgTrpValGlnGlnAsnIleAlaHisPheGlyValAsnProAspArg 239
Oy 771 GTGACCATCTTGTGGAGAGTCCGCGAGAGCCATAGTGTCTTATGCTTATGCTTCC 830
Db 240 ValThrIlePheGlyGlySerIleGlyIleThrSerValSerLeuValValSerPro 259
Oy 831 ATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGCTGCTCATCCCTTAC 890
Db 260 IleSerGlnGlyLeuPheHisGlyAlaIleMetGluSerGlyValAlaLeuLeuProGly 279
Oy 891 CTGAGAGCCCATGATTAGAGAGAGTGAAGCTG-----CAGTGTGTGCAATTTC 944
Db 280 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaHisLeu 295
Oy 945 TGTGGTAACTAATGCGTCAAGTCTGAGGCTTGTGAGTGTGCTGAGGACAAACCTTCC 1004
Db 296 SerAlaCysAspGlnValAspSerGlnAlaLeuValGlyCysLeuArgGlyIysSerIys 315
Oy 1005 AAGAGTGTGAGCCCTGAGCCAGCAAAACAAAGTCTTCACTGAGTGTGAGTGTGCT 1064
Db 316 GluGluIleLeuAlaIleAsnIleProPheIlePheGlyValValAlaAspGlyVal 335
Oy 1065 TTCTTTCTTAATGAGCTCTAGATCTATTTGTCTCAAAAAGCTTAAAGCAATTCTTCC 1124
Db 336 PheLeuProArgHisAspGlnGlnIleuLeuAlaSerAlaAspPheGlnProValProSer 355
Oy 1125 ATCATCGAGTCAATTAACACAGAGTGTGCTTCTGCTGCT-----ATGAAG 1172
Db 356 IleValGlyValAsnAsnAsnGlnPheGlyTrpLeuIleProIysValMetArgIleTyr 375
Oy 1173 GAGGCTCTGAGATCTTCAAGTGTGCTCAACAAAGCTTGTCCCTTCACTGTATCAAAAC 1232
Db 376 AspThrGlnIleGlyMetAspArgGlnAlaSerGlnAlaIleGlnIleMetLeuThr 395

```

```

Oy 1233 ATCTGCATATCCCGCTCAAGTATTGCACTTGTGGTAAATGAATCTTCATGACAAG 1292
Db 396 LeuLeuMetLeuProProThrPheGlyAspLeuValArgGlnGluTyrIleGlyAspAsn 415
Oy 1293 CACTCCCTGACTGAAATCCGAGACAGTCTTCTTGACATTCCTTGGAGATGTGTTTGG 1352
Db 416 GlyAspProGlnThrLeuGlnAlaGlnPheGlnGlnMetMetAlaAspSerMetPheVal 435
Oy 1353 GTCCCTGACTGATATCAAGCTGATATCAAGAGATGTGGTGGTCACTGTACTTCTAT 1412
Db 436 IleProAlaLeuGlnValAlaHisPhe---GlnCysSerArgAlaProValIleTyrPheTyr 454
Oy 1413 GAGTTTCCGACACCGGCTCAGTGTCTTGAAGACAGAACCGGCTTTTGTCAAGCCGAC 1472
Db 455 GluPheGlnHisGlnProSerTrpLeuIysAsnIleArgProProHisMetCysValAsp 474
Oy 1473 CACGCTGATGAATCCGCTTGTGTGCTGTGCTGCTTCTCAAGGGGACATTTGTATG 1532
Db 475 HisGlyAspGluLeuProPheValPhe--ArgSerPhePheGlyGlyAsnTyrIleIys 493
Oy 1533 TTGGAAGAGCCACGAGAGAGAGTAACTGAGCCGGAAGATGATGAATATCTGGGCT 1592
Db 494 Phe-----ThrGlnGlnGlnIleGlnIleuSerArgIleMetCysTyrTrpAla 510
Oy 1593 ACCTTTGTGCAACCGGAGATCTTATGGAAGCACTGTCTGTGTGCGCACTTATAT 1652
Db 511 AsnPheAlaArgAsnGlyAsnProAsnGlyGlnGlyLeuProHisIleTrpProLeuAsp 530
Oy 1653 CTGACTGAGACATCTTCAAGTGTGAATGAGCTTGGACAGAGACTCAAGAA 1712
Db 531 GlnGlnGlnIleGlyLeuGlnIleuAsnLeuGlnProAlaValGlyArgAlaLeuIysAla 550
Oy 1713 CCGCGGTGAGTATTGGAACAGACATCCCC 1745
Db 551 HisArgLeuGlnPheTrpIleValAlaLeuPro 561

```

RESULT 5

```

US-09-595-682B-21
; Sequence 21, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danke, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 21
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-595-682B-21

```

```

Alignment Scores:
Pred. No.: 1,67e-103 Length: 565
Score: 1150.50 Matches: 252
Percent Similarity: 56.97% Conservative: 75
Best Local Similarity: 43.90% Mismatches: 180
Query Match: 29.71% Indels: 67
DB: 4 Gaps: 13

```

```

US-10-023-515-1 (1-2158) x US-09-595-682B-21 (1-565)
Oy 129 TGGTCTTTTCTGATTTCTCAGCCCTGTGGAGACAGACAGTGGGAAAAAATCGG 188
Db 129 TGGTCTTTTCTGATTTCTCAGCCCTGTGGAGACAGACAGTGGGAAAAAATCGG 188

```

```

Db      2 TrpLeuCyAlaLeuAlaLeuAlaSerLeuAlaAlaCysThrAlaTrpGlyHis----- 19
QY      189 CTTTCCTGGTGAAGGGCCACAGAGAAACCAAGCTGGAGATTCAGGCAAGCAAGCTC 248
Db      20 ProSerAlaPro--ProValAlaThrThrValHisGlyLeuValLeuGlyLeuPheVal 38
QY      249 ACTGTGCTGGGAAGCCCTGTGCTGGTGAAGCTGTTCCTGAGATGCCCTTGGCTGCTCC 308
Db      39 SerLeuGluGlyPheAlaGlnProValAlaValPheLeuGlyValProPheAlaLeuPro 58
QY      309 CCGCTGGGATCCCTGCGGATTTACGAACCCGAGCTGCATGCGCTGGAGTAATCTGCGA 368
Db      59 ProLeuGlySerLeuAlaGlyPheAlaProGlnProAlaGlnSerTrpSerHisValLys 78
QY      369 GAAGCACTCTCAATTTGTGGCTCCAGAACTCAAGTGGCTGCTTGAATCA 428
Db      79 AanthrThrSerTyrrProPheMetCysSerGlnAspAla-----ValSerGly 94
QY      429 CACATGCTC-----AAGGTGATTTACCGG--AAATTGCGAGTG 464
Db      95 HisMetLeuSerGlyLeuPheThrAsnAlaGlySerGlnMetLeuPheProLeuLysPhe----- 112
QY      465 TCAGAAAGATGCGCTCACTGCAATATGATGATGCGCTGCCACGCGGATACAGGCTCCAG 524
Db      113 SerGlnAspCysLeuTyrrLeuAlaMetLeuTyrrThrProAlaAspLeuThrLysAlaGlyAsp 132
QY      525 CTCCCGCTTGTGTGTGTCCAGAGAGTGCCTTCAAGACTGGCTCAGGCTGCATCTTT 584
Db      133 LeuProValMetValTrpIleHisGlyGlyGlyLeuMetValGlyGlyAlaSerThrTyrr 152
QY      585 GATGGGCTCCGCGCTGCTGCTGATGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 644
Db      153 AspGlyLeuAlaLeuSerAlaHisGlnAsnValValValValThrIleGlnTyrrAspLeu 172
QY      645 GGAATATTGGTGTCTTCAACACATGAGATCAGATGCTCCGCGGAGCTGGCTTCAAG 704
Db      173 GlyIleTrpGlyPhePheSerThrGlyAspGlnHisSerAlaGlyValAsnTrpIleHisLeu 192
QY      705 GACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
Db      193 AspGlnValAlaAlaLeuAlaGlyTrpValGlnAspAsnIleAlaAsnPheGlyGlyAspPro 212
QY      765 AGCTGTGTGACCATCTTGTGGCGAGTCCGCGGAGGACATAGTGTTCATGTTTATGCT 824
Db      213 GlySerValThrIlePheGlyLeuSerAlaGlyGlyGlnSerValSerIleLeuLeuLeu 232
QY      825 TCTCCATGAGCCAAAGCTTATTCACAAAGCATATGAGAGTGGGAGTGGCCATCATC 884
Db      233 SerProLeuThrLysAsnLeuPheHisArgAlaIleSerGlnSerGlyValAlaLeuLeu 252
QY      885 CCTTACCTGGAGGCCCATGATTAAGAAAGAGTGAAGAC-----CTCCAGGTGGTTGCA 938
Db      253 SerSerLeu-----PheArgLysAsnThrLysSerLeuAlaGlnIleAla 268
QY      939 CATTTCTGTGTAAATGCGTCAAGACTCTGAGCGCTGCTGAGTGGCTGAGGACAA 998
Db      269 IleGlnAlaGlyCysLeuThrThrSerAlaValMetValHisCysLeuAlaGlnLys 288
QY      999 CCTTCAAGAGAGTGGCTG-----ACCTTCAAGCAAGAA----- 1031
Db      289 ThrGlnGlnGlnLeuMetGlnValThrLeuLysMetLysPheMetAlaLeuAspLeuVal 308
QY      1032 -----ACAAAGCTTTTCACTGAGTGGTGAATGAGTGGCTTTCTTCTT 1073
Db      309 GlyAspProLysGlnAsnThrAlaPheLeuThrThrValIleAspGlyValLeuLeuPro 328
QY      1074 AATGAGCTTGAATCTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGGA 1133
Db      329 LysAlaProAlaGlnIleLeuAlaGlnLysLeuTyrrAsnMetLeuProTyrrMetValGly 348
QY      1134 GTCAATTAACCAAGAGTGTGGCTTCTGCTGCTGCTTATG----- 1169
Db      349 IleAsnGlnGlnGlnPheGlyTrpIleIleProMetGlnMetLeuGlyTyrrProLeuSer 368

```

```

QY      1170 -----AAGAGGCTCTGAGATCTTCAGTGGCTCCAAAGTCCCTT 1211
Db      369 GlnGlyLysLeuAspGlnLysThrAlaThrGlnLeuLeuTrpLysSer----- 384
QY      1212 GCCCTCATCTGATACAAACATCCGACATCCCGCTCAGATATTTGACCTTGGACT 1271
Db      385 -----TyrrProIleValAsnValSerLysGlnLeuThrProValAlaThr 399
QY      1272 AATGAATATTCTTCATGACAGACATCCCTGACTGAATCCGAGACAGTCTTGGACTTG 1331
Db      400 GlnLysTyrrLeuGlyGlyThrAspAspProValLysLysLysAspLeuPheLeuAspMet 419
QY      1332 CTGGAGATGATGTTCTTTGTGGTCCCTGACATGATCAGCTGCAATATCAGAGATGCT 1391
Db      420 LeuAlaAspLeuPhePheGlyValProSerValAsnValAlaAspHisHisArgAspAla 439
QY      1392 GGTGCACTGTCTTACTTATGATGTTTGGCAGCCGCTCAGTGCCTTGAAGACAGAG 1451
Db      440 GlyAlaProThrTyrrMetTyrrGlyIleTyrrArgTyrrArgProSerPheSerAspMetArg 459
QY      1452 CCGGCTTTTTCAAAGCGGACGACATGAAAGTCCGCTTTGTGGTGGTGGCTTC 1511
Db      460 ProLysThrValIleGlyAspHisGlyAspGlnIlePheSerValLeuGlyAlaProPhe 479
QY      1512 CTGAAGGGGAGCATGTGTTATGTTTGAAGAGCCACGAGAGAGAGAAATTACTGAGCCGG 1571
Db      480 LeuLys-----GlnGlyAlaThrGlnGlnGlnIleLysLeuSerLys 493
QY      1572 AAGATGATGAATATCTGGGCTACTTTGCTGGAACCGGAAATCTTAATGGAACGACTTG 1631
Db      494 MetValMetCysTyrrTrpAlaAsnPheAlaArgAsnGlyAsnProAsnGlnGlyLeu 513
QY      1632 TCTGTGGGCAAGTTAATATCTGACTGAGCAGTACTGCTCAGTGCAGTGAACATGAGAC 1691
Db      514 ProGlnThrProAlaTyrrAspTyrrLysGlnGlyTyrrLeuGlnIleGlyAlaThrGln 533
QY      1692 CTGCAAGAGACTCAAAAGAACCGCGGTGGATTTTGGACC 1733
Db      534 AlaAlaGlnLysLeuLysAspLysGlnValAlaAlaPheTrpThr 547

RESULT 6
US-09-264-737-2
/ Sequence 2, Application US/09264737A
/ Patent No. 6107549
/ GENERAL INFORMATION:
/ APPLICANT: Feng, Paul C.C.
/ APPLICANT: Ruff, Thomas G.
/ TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
/ FILE REFERENCE: 38-21(10551) R1R3 Pyridine Tolerance
/ CURRENT FILING DATE: US/09/264,737A
/ EARLIER APPLICATION NUMBER: 60/077,377
/ EARLIER FILING DATE: 1998-03-10
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 566
/ TYPE: prt
/ ORGANISM: Rabbit
US-09-264-737-2

Alignment Scores:
Pred. No.: 2,5e-102
Score: 1138,50
Percent Similarity: 56,62%
Best Local Similarity: 43,55%
Query Match: 29,40%
DB: 3
Gaps: 14

US-10-023-515-1 (1-2158) x US-09-264-737-2 (1-566)

```


Query Match: 29.20% Indels: 67
 DB: 4 Gaps: 13
 US-10-023-515-1 (1-2158) x US-09-595-682B-26 (1-543)

QY 129 TGGTCTTTTTCGATTTCTCAGCCCTGTGGGACACAGACAGTGGGAAAACTGGG 188
 DB 2 TTPLeuCyValAlaValAlaLeuAlaSerLeuAlaAlaCysThrAlaTrpGlyHis----- 19
 QY 189 CCTTCTGTGAAGGGCCACAGAGGAACCCAGCGCTGGGATTCAGGCGCAAGATC 248
 DB 20 ProSerAlaPro-----ProValValAspThrValHisGlyValValLeuGlyLysPheVal 38
 QY 249 ACTGTCTGGGAAGCCCTGTGCTGTGACAGTGTCTCCGAGAGTCCCTTGTGCTGCC 308
 DB 39 SerLeuGlyGlyPheAlaGlnProValAlaValAlaPheLeuGlyValProPheAlaLysPro 58
 QY 309 CCGCTGGGATCCCGGACATTTACGAACCCGACCTGCATCGCCCTGGATTAATTGGCA 368
 DB 59 ProLeuGlySerLeuAlaGlyPheAlaProProGlnProAlaGlnSerTrpSerHisValLys 78
 QY 369 GAAGCCACCTCTACCTTAATTTGTGCTCCAGAACCTCAGAGTGGCTCTTACATCA 428
 DB 79 AsnThrTrpSerTrpProPheCysSerGlnAspAla-----ValSerGly 94
 QY 429 CACATGCTC-----AAGGTGATTAACCG-----AAATTCGAGATG 464
 DB 95 HisMetLeuSerGlnLeuPheThrAsnAlaGlyGlnValAsnLeuProLeuLysPhe----- 112
 QY 465 TCAGAAAGACTGCTTACCTGAAACATCTTATGCGGCTCCACGCCATACAGGCTCCAA 524
 DB 113 SerGlnAspCysLeuTrpLeuAsnHisLeuTrpAlaAspLeuThrLysAlaGlyLys 132
 QY 525 CTCCCGCTGTGGTGTGTCCAGAGAGTGCCTTACAGACTGGCTCAGCTCCATCTTT 584
 DB 133 LeuProValMetValTrpHisGlyGlyGlyLeuMetValGlyLysAspThrTrp 152
 QY 585 GATGGGTCCGCGCTGCTGCTATGAGACGTCGCTGTGTGTGTGTGTGTGTGTGTGTGT 644
 DB 153 AspGlyLeuAlaLeuSerAlaHisGlnAsnValValValAlaThrHisGlnLysArgLeu 172
 QY 645 GGAATATTTGGTTTCTTCAACATAGGATGAGCATGTCCCGGAACTGGGCTTCAAG 704
 DB 173 GlyIleTrpGlyPhePheSerThrGlyAspGlnHisSerArgGlyAsnTrpGlyHisLeu 192
 QY 705 GACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
 DB 193 AspGlnValAlaAlaLeuAlaGlyTrpValGlnAspAsnHisAlaAsnPheGlyGlyAspPro 212
 QY 765 AGCTGTGACCATCTTTGGCGAGTCCGCGGAGCATTAAGTGTTCAGTCTTACTG 824
 DB 213 GlySerValThrHisPheGlyGlnSerHisGlyGlnSerValSerHisLeuLeuLeu 232
 QY 825 TCTCCCATGGCCAAAGGCTTATTCACAAAGCATGAGAGTGGGCTGGCCATCATC 884
 DB 233 SerProLeuThrLysAsnLeuPheHisArgAlaHisSerGlnSerGlyValAlaLeuLeu 252
 QY 885 CCTTACCTGAGGAGCCATGATTAAGAGAGTGAAGC-----CTGCAGGTGGTTCGA 938
 DB 253 SerSerLeu-----PheArgLysAsnThrLysSerLeuAlaGlyLysHisAla 268
 QY 939 CATTTCTGTGTAAACAATGCTCAGACTCTGAGCCCTGCTGAGAGTCCGAGGACAA 998
 DB 269 HisGlnAlaGlyCysLysThrThrSerHisValMetValHisCysLeuAlaGlnLys 288
 QY 999 CCTTCAAGAGAGCTGCTG-----ACCTGACCAAGAA----- 1031
 DB 289 ThrGlnGlnGlnLeuMetGlnValThrLeuLysMetLysPheMetAlaLeuAspLeuVal 308
 QY 1032 -----ACAAAGCTTTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1073
 DB 309 GlyAspProLysGlnAsnThrAlaPheLeuThrThrValHisAspGlyValLeuLeuPro 328

QY 1074 AATGACCTCTAGATTAATGTCTCAGAAAGCATTTAAAGCAATTCTTCATCATCGGA 1133
 DB 329 LysAlaProAlaGlnHisLeuAlaGlnLysLysTrpAsnMetLeuProTrpMetValGly 348
 QY 1134 GTCAATTAACCAAGAGTGGCTCTTCTGCTGCTATG----- 1169
 DB 349 HisGlnGlnGlnLysPheGlyTrpIleIleProMetGlnMetLeuGlyTrpProLeuSer 368
 QY 1170 -----AAGAGGCTCTGAGATTCCTCACTGAGTCTTCAACAACTCCCTT 1211
 DB 369 GlnGlyLysLeuAspGlnLysThrAlaThrGlnLeuLeuTrpLysSer----- 384
 QY 1212 GCCCTTCACTTGATTAACAAATCTTCACATCCCGCTCAGATTTTGTGCACTTGTGCT 1271
 DB 385 -----TrpProIleValHisValSerLysGlnLeuThrProValAlaThr 399
 QY 1272 AATGAATCTTCCATGACAAACACACTCCCTGATCTGAAATCCGAGACACTTCTTGACT 1331
 DB 400 GlnLysTrpLeuGlyGlyThrAspAspProValLysLysAspLeuPheLeuAspMet 419
 QY 1332 CTGGAAGATGTCTTCTTGTGTGCTCCCTGACATGATCAGACTGATATCAGAGATGCT 1391
 DB 420 LeuAlaAspLeuLeuPheGlyValProSerValAsnValAlaArgHisHisAspAspAla 439
 QY 1392 GATGACCTGTCTACTTCTTATGACTTTCGACACCGGCTTCAAGCTTTTGAAGACAGAA 1451
 DB 440 GlyAlaProThrTrpMetTrpGlyTrpArgTrpArgProSerPheSerAspMetArg 459
 QY 1452 CCGGCTTTTGTCAAGCCGACACAGCTGAGATGAGTCCGCTTGTGTGCTGCTGCTTC 1511
 DB 460 ProLysThrValHisGlyAspHisGlyAspGlnHisPheSerValLeuGlyAlaProPhe 479
 QY 1512 CTGAAGGGGACATTTGATGTTTTCGAAGAGACCAAGAGAGAGAGAACTTACTGAGCCG 1571
 DB 480 LeuLys-----GlnGlyAlaThrGlnGlnGlnLysLeuSerLys 493
 QY 1572 AAGATGATGAATATCTGGGCTACTTGTGCTCGAACCGGGAATCTTAATGGAAACACCTG 1631
 DB 494 MetValMetLysTrpTrpAlaAsnPheAlaAspAsnGlyAspProAsnHisGlnLysLeu 513
 QY 1632 TCTCTGTGGCCAGCTTATTAATCTGACTGAGCAGTACTTCCAGCTGACTTGAATGAGC 1691
 DB 514 ProGlnTrpProAlaTrpAspTrpLysGlnGlyTrpLeuGlnHisGlyAlaThrThrGln 533
 QY 1692 CTCGACAGAGACTCAAGAACCGCGGCTG 1721
 DB 534 AlaAlaGlnLysLeuLysAspLysGlnVal 543

RESULT 8
 US-08-845-295A-2
 / Sequence 2, Application US/0845295A
 / Patent No. 5817490
 / GENERAL INFORMATION:
 / APPLICANT: Hubbe, John C.
 / TITLE OF INVENTION: Enzymatic Process for the Manufacture of
 / TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
 / NUMBER OF SEQUENCES: 3
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Eastman Chemical Company
 / STREET: P.O. Box 511
 / CITY: Kingsport
 / STATE: Tennessee
 / COUNTRY: USA
 / ZIP: 37662-5075
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: 3.5 inch disk
 / COMPUTER: IBM Compatible
 / OPERATING SYSTEM: Windows 95
 / SOFTWARE: Microsoft Word
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/845,295A
 / FILING DATE: 25-Apr-11-97

CLASSIFICATION: 435
 PRIOR APPLICATION DATA: US 60/017,879
 APPLICATION NUMBER: 17-MAY-1996
 FILING DATE: 17-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Cheryl J. Tubach
 REGISTRATION NUMBER: 38,346
 REFERENCE/DOCKET NUMBER: 70432
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 423-229-6189
 TELEFAX: 423-229-1239
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 584 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 US-08-845-295A-2

Alignment Scores:
 Pred. No.: 1,02e-97 Length: 584
 Score: 1091.50 Matches: 259
 Percent Similarity: 54.39% Conservative: 63
 Best Local Similarity: 43.75% Mismatches: 191
 Query Match: 28.19% Indels: 79
 DB: 2 Gaps: 17

US-10-023-515-1 (1-2158) x US-08-845-295A-2 (1-584)

QY 129 TGGTCTTTTCTGATATCTCAGCCCTGTGGGACAGACATGGGAAAACTGGG 188
 DB 2 TrrleuLeuProleuValLeuThrSerleuAlaSerleuAlaThrTrp-----AlaGly 19
 QY 189 CTTTCTGGAAGGCGCACAGAGAACACAGGCTGGGATTCAGGCGCAAGCTAC 248
 DB 20 GlnProAlaSerProValValAspThrAlaGlnGlyArgValLeuGlyIystrVal 39
 QY 249 ACTGTG-----CTGGGAAGCCCTGTGCTGAACGTGTGCTGGAGTCCCTTTGCT 302
 DB 40 SerleuGlnGlyLeuAlaPheThrGlnProValAlaValPheLeuGlyAlaProPheAla 59
 QY 303 GCTCCCGCGGTGGATCCCTGGGATTTACGAACCGGACGCTGCATCGCTGGGATTAAC 362
 DB 60 LysProProleuGlySerleuArgPheAlaProProGlnProAlaGlnIuproTrpSerPhe 79
 QY 363 TTGCGAAGACCACTCTTACCTTAATTGTGCTTGCAGAACTCA-----GAGTGGCTG 416
 DB 80 ValLysAsnThrThrSerTyxProPrometCysCysGlnAspProValValGlnGlnMet 99
 QY 417 CTCTTAAGT-----CAACACATGCTCAAGGTGCATTACCCGAAA 455
 DB 100 ThrSerAspLeuPheThrAsnPheThrGlyLysGlnLysLeuThrLeuGlnPhe----- 117
 QY 456 TTCGAGTGTCAAGAGACTGCTCTACCTGAACATCTATGCGCTGCCACCGCGATACA 515
 DB 118 -----SerGlnAspCysLeuTyxLeuAsnIleTyxThrProAlaAspLeuThrLys 134
 QY 516 GCGTCCAAAGCTCCCGTCTGGTGTGTTCCAGAGAGTGCCTTCAAGATGGCTCAAGCC 575
 DB 135 ArgGlyArgLeuProValMetValTrpIleAspGlyGlyLeuValLeuGlyAla 154
 QY 576 TCCATCTTTGAGGGTCCGCGCTGGCTGCTATGAGAAC-----GTGCTGGTGTGGTC 629
 DB 155 PrometTyxAspGlyValValLeuAlaAlaAspGlnAspPheThrValValValAla 174
 QY 630 GTCCAGTACCGGCTGAATATTTGTTTCTTCAACATGGAGATCAAGATGCTCCGGGG 689
 DB 175 IleGlnTyxArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGlnIleSerArgGly 194
 QY 690 AACTGGGCTTCAAGACCAAGGTGGCTGTCTGTCTGGGCTCAAGAAACATCGAGTTTC 749
 DB 195 AsnTrpGlnIleLeuAspGlnValAlaAlaLeuAspIleTrpValGlnGlnLeuAlaAsn 214

QY 750 TTCGTGGGGAACCCAGACTCTGTGACCATTTTGGCGAGTCC-----GCGGAGCCATA 803
 DB 215 PheGlyGlyAspProGlySerValThrIlePheGlyGlySerPheThrAlaGlyGly 234
 QY 804 AGTGTTCATGCTTAACTGTCTCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATG 863
 DB 235 SerValSerValLeuValLeuSerProleuAlaLysAsnLeuPheIleArgAlaIleSer 254
 QY 864 GAGAGTGGGGTGGCCATCCCTTACCTGAGAGGCCCATGATTAAGTAAGAGTGAAGC 923
 DB 255 GluSerGlyValAlaLeuThrValAlaLeuValArgLysAspMetCysAlaAlaLys 274
 QY 924 CTCAGTGGTTCACATTTCTGTGTGTAACAATGCGTCAAGCTGAGGCGCTGAGAG 983
 DB 275 GlnIleAlaValLeuAlaGlyCysLysThrThrThrIleSerAlaValPheThrPheValHis 294
 QY 984 TGCCTGAGGACAAACCTCCCAAGAGCTGTGACCTTCAGCCGAGAAAACAAAGCTTTTC 1043
 DB 295 CysLeuArgGlnLysSerGlyAspGlyLeuLeuAspLeuThrLeuLysMetLysPheLeu 314
 QY 1044 ACT-----CGAGTGGTGTAT 1058
 DB 315 ThrLeuAspPheHisGlyAspGlnArgGlnLysSerHisProPheLeuProThrValValAsp 334
 QY 1059 GGTGCTTTCTTTCTTAATGAGCTCTAGATCTATTGTCTCAGAAA-----GCATTTAAA 1112
 DB 335 GlyValLeuLeuProLysMetProGlnGlnIleLeuAlaGlnLysAspPheThrPheAsn 354
 QY 1113 GCAAATTCCTTCATCATCGAGATCAATTAACAAGAGTGGCTTCTGCTGCT----- 1166
 DB 355 ThrValProTyxIleValGlyIleAsnLysGlnGlnPheGlyTrpLeuLeuProThrMet 374
 QY 1167 -----ATGAAGAGCTTCCTAGATCCCTCAT 1193
 DB 375 MetGlyPheProLeuSerGlnGlyLysLeuAspGlnLysThrAlaThrSerLeuLeuTrp 394
 QY 1194 GCGTCCAAAGTCCCTTCCCTCCATCGATGATCAAAACATCTGCACATCCCGCTCAG 1253
 DB 395 LysSer-----TyrProIleAlaAsnIle-----ProGlu 404
 QY 1254 TATTTGACCTTGTGCTAATGATATCTTCATGACAAAGC-----TCC 1298
 DB 405 GluLeuThrProValAlaThr-----PheThrAspLysTyxLeuGlyGlyThrAspAsp 422
 QY 1299 CTGACTGAATCCGAGACAGTCTTGTGCACTTGTGCAAGATGTGTCTTGTGCTGCT 1358
 DB 423 ProValLysLysLeuAspLeuPheLeuAspLeuMetGlyAspValValPheGlyValPro 442
 QY 1359 GCAGTATCAACAGCTCGATATGACAGAGATGCTGTGCACTGTCTACTTATGAGTTT 1418
 DB 443 SerValThrValAlaArgGlnHisArgAspAlaGlyAlaProGlyThrTyxMetTyxIlePhe 462
 QY 1419 CGGACCGGCTCAAGTGTCTTGAGAC-----ACGAACCGGCTTTGTCAAGCCGAC 1472
 DB 463 GlnTyxArgProSerPheSerSerAspLysPheThrLysProLysThrValIleGlyAsp 482
 QY 1473 CACGCTGATGAAGTCCGCTTGTGCTGCTGCTGCTTCTGAAAGGGACATTTGATG 1532
 DB 483 HisGlyAspGlnIlePheSerValPheGlyPheProLeuLeuLysGlyAsp----- 499
 QY 1533 TTCGAAGGACCAAGGAGGAGAAAGTATCAAGACCGGAGATGATGAATAAGTGGGT 1592
 DB 500 -----AlaProGlnGlnGlnValSerLeuSerLysThrValMetLysPheThrAla 516
 QY 1593 AACTTGTCTGAACCGGGAATCTTAATGGAACGACTGTCTGTGGCA-----GCT 1646
 DB 517 AsnPheAlaArgSerGlyAsnProAsnGlyGlyLeuProHisIleTrpProPheThrMet 536
 QY 1647 TATTAATGATGACAGACTCTCAAGCTGCACTTGAACATGAGCTGGAGACAGACTTC 1706
 DB 537 TyxAspGlnGlnGlnGlyTyxIleGlnIleGlyValAsnThrGlnAlaAlaLysArgLeu 556
 QY 1707 AAAGAACCGGCGGTGATTTTTCAGACAGACATTC 1742

```

Db      557  LysGlyGluGluValAlaPheThrPheAspLeu 568
RESULT 9
/ Sequence 2, Application US/09140933
/ Patent No. 6022719
/ GENERAL INFORMATION:
/ APPLICANT: Hubbs, John C.
/ TITLE OF INVENTION: Enzymatic Process for the Manufacture of
/ TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-Gulonic Acid, and Esters of
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Eastman Chemical Company
/ STREET: P. O. Box 511
/ CITY: Kingsport
/ STATE: Tennessee
/ COUNTRY: USA
/ ZIP: 37662-5075
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 inch disk
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: Windows 95
/ SOFTWARE: Microsoft Word
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/140,933
/ FILING DATE: 27-August-98
/ CLASSIFICATION:
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: US 60/017,879; 08/845,295
/ FILING DATE: 17-May-96; 25-April-97
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Cheryl J. Tudach
/ REGISTRATION NUMBER: 38,346
/ REFERENCE/DOCKET NUMBER: 70432
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 423-229-6189
/ TELEFAX: 423-228-1239
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 584 amino acids
/ TYPE: Amino Acid
/ TOPOLOGY: Linear
/ MOLECULE TYPE: protein
/
US-10-023-515-1 (1-2158) x US-09-140-933-2 (1-584)
Alignment Scores:
Pred. No.: 1.02e-97 Length: 584
Score: 1091.50 Matches: 259
Percent Similarity: 54.39% Conservative: 63
Best Local Similarity: 43.75% Mismatches: 191
Query Match: 28.19% Indels: 79
Gaps: 17

```

```

QY      363  TTGCGAAGACCACTCTACCTAATTGTCCTCGAAGTCA-----GAGTGGCTG 416
Db      80  ValIyAsnThrThSerYrProPheMetCysGlnAspProValValGlnIleMet 99
QY      417  CTCTTGAT-----CAACATGCTTCAGTGCATTACCCGAAA 455
Db      100  ThrSerAspLeuPheThrAsnPheThrGlySerGluArgLeuThrLeuGlnPhe----- 117
QY      456  TTGCGAGTGTGAGAAGAGCTCTACCTACCAATCATGTCGCGCTCCGACCGCATCA 515
Db      118  -----SerGlnAspCysLeuYrLeuAsnIleYrThrProAlaAspLeuThrLys 134
QY      516  GGCTCCAGCTCCCGCTGTGTGTGTGTGTCCAGAGGTGCTTCAGACTGCTCAGCC 575
Db      135  ArgGlyArgLeuProValMetValTrpIleHisGlyGlyLeuValLeuGlyValAla 154
QY      576  TCCATCTTGATGGGTCCGCGCTGCTGCTCCATAGAC-----GTGCTGTGTGCTC 629
Db      155  ProMetYrAspGlyValValLeuAlaHisGlnAsnPheThrValValAla 174
QY      630  GTCCAGTACCGGCTGGAATATTTGTTTTCACACATGAGTACGATCGCCGGGG 689
Db      175  IleGlnYrArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGlnHisSerArgGly 194
QY      690  AACTGGGCTTCAGAGACACAGTGGCTGCTGTCTGCTGGGTCCAGAAACATCGACTTC 749
Db      195  AsnTrpGlyHisLeuAspGlnValAlaAlaLeuHisTrpValGlnIleAsnIleAlaAsn 214
QY      750  TTGCGTGGGACCCCACTCTGTGACCATCTTTGGCAGATC-----GGGGAGCCATA 803
Db      215  PheGlyIyAspProGlySerValThrIlePheGlyGlnSerPheThrAlaGlyGlyGln 234
QY      804  AGTGTTCCTACTCTTAATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATCATG 863
Db      235  SerValSerValLeuValLeuSerProLeuAlaIyAsnLeuPheHisArgAlaIleSer 254
QY      864  GAGAGTGGGTGGCCATCATCCCTTACCTGGAGCCCATGATATAGAAAGCTGAGAC 923
Db      255  GluSerGlyValAlaLeuThrValAlaLeuValArgLysAspMetLysAlaAlaLys 274
QY      924  CTGCAAGTGGTTCACATTTCTGTGTAAACATGCGCTGAGCTTGAGCCCTGTCAGAG 983
Db      275  GlnIleAlaValLeuAlaGlyCysLysThrThrThrSerAlaValPheThrPheValHis 294
QY      984  TGCCTGAGCAAAACCCCTCCAGAGCTGCTGACCTCAGCCGAAACAAAGTCTTTC 1043
Db      295  CysLeuArgGlnLysSerGlnAspGlnLeuLeuAspLeuThrLeuLysMetLysPheLeu 314
QY      1044  ACT-----CGAGTGGTGTAT 1058
Db      315  ThrLeuAspPheHisGlyAspGlnArgGlnLysSerHisProPheLeuProThrValAlaAsp 334
QY      1059  GGTGCTTCTCTTCCATAATAGCTCTAGATCTATTGTCTCAGAAA-----GCATTTAAA 1112
Db      335  GlyValLeuLeuProLysMetProGlnGlnIleLeuAlaGlnLysAspPheThrPheAsn 354
QY      1113  GCAATCTCTTCATCGAGTCAATTAACAAGAGTGGCTCTCTGCTCTCTCTCTCTCT 1166
Db      355  ThrValProYrIleValGlyIleAsnLysGlnGlnLysPheGlyTrpLeuLeuProThrMet 374
QY      1167  -----ATGAAGAGGCTCTGAGATCTCTCAGT 1193
Db      375  MetGlyPheProLeuSerGlnLysLysLeuAspGlnLysThrAlaThrSerLeuLeuTrp 394
QY      1194  GGCTCCACAAGTCCCTTGCCTCATGATGATACAAACATCTGACATCTCCGCTCAG 1253
Db      395  LysSer-----TyrProIleAlaAsnIle-----ProGln 404
QY      1254  TATTTCACCTTGTGGCTAATGAATATCTTCATGACAAAGAC-----TCC 1298
Db      405  GlnLeuThrProValAlaThr-----PheThrAspYrLysLeuGlyGlyThrAspAsp 422
QY      1299  CTGACTGAATATCCAGACAGTCTTCTGACATTGCTTGAGATGTGTTTGTGGTCTCT 1358

```



```

Db      423 ProvallylelyserleupheleuAapleuMecglyAepValPheglyValPro 442
Qy      1359 GCACTGATATACAGCTGATATTCACAGAGATGCTGGTGCACCTGTCTACTTATGAGTTT 1418
Db      443 SerValThrValAlaIArgInIAsrGAspAlaGlyAlaIaProThrIleuTyrgluPhe 462
Qy      1419 CGGACCGGCTCTGAGTGTCTTGAAGAC-----ACGAAGCGGCTTTTGTTCGAAGCGGAC 1472
Db      463 GlnTyArgProSerPheSerSerAspIlyPheThrIlyPheProIlySerThrValIleGlyAsp 482
Qy      1473 CACGCTGATGAATCCGCTTTGTGTTCGCTGCTGCTTCTCGAAGGAGGACATGTTATG 1532
Db      483 HLeGlyAspGluIlePheSerValPheGlyPheProIleuIleuIleGlyAsp----- 499
Qy      1533 TTCGAAGAGCCACGAGAGAGAGAAATTACTAGCCCGGAGATGATGAATATCTGGCT 1592
Db      500 -----AlaProIleuGluIleuValSerIleuSerIleuThrValIleuIlySerPheTrpAla 516
Qy      1593 ACCTTGTCTGAACCGGAGATCCTATGGGAACGACTGTCTGTGGCCA-----GGT 1646
Db      517 AsnPheAlaArgSerIlyAsnProAsnGlyGluIleuProHisTrpPheThrMet 536
Qy      1647 TATATCTGACTGAGAGTACCTCTGAGCTGAGTGAACATGAGCCTCGGACAGAGACTC 1706
Db      537 TyrAspGluGluGluIleuTyrgluGluIleGlyValAsnThrGlnAlaIleuIlyAsrGlu 556
Qy      1707 AAGAAGCCGCGGTGATTTTGGACCAGACCATC 1742
Db      557 LysGlyGluIleuValaIaPheTrpAsnAspLeuIleu 568

```

RESULT 10

US-09-146-661-2
 ; Sequence 2, Application US/09146661
 ; Patent No. 6136575

```

; GENERAL INFORMATION:
; APPLICANT: Hubbs, John C.
; TITLE OF INVENTION: Enzymatic Process for the Manufacture of
; TITLE OF INVENTION: Ascorbic Acid, 2-Keto-L-gulonic Acid, and Esters of 2-Keto-L-Gu
; TITLE OF INVENTION: Acid
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Eastman Chemical Company
; STREET: P.O. Box 511
; CITY: Kingsport
; STATE: Tennessee
; COUNTRY: USA
; ZIP: 37662-5075
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,661
; FILING DATE: 03-September-98
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,879; 08/845,295
; FILING DATE: 17-May-96; 25-April-97
; ATTORNEY/AGENT INFORMATION:
; NAME: Cheryl J. Thibach
; REGISTRATION NUMBER: 38,346
; REFERENCE/DOCKET NUMBER: 70432
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 423-229-6189
; TELEFAX: 423-229-1239
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 584 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein

```

US-09-146-661-2

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1,02e-97	1091.50	54.39%	43.75%	26.19%	584	259	63	191	79	17

US-10-023-515-1 (1-2158) x US-09-146-661-2 (1-584)

```

Qy      129 TGGTGTCTTTTCTGATCTCCAGCCCTGTTGGACACAGACAGTGGGGAATACTGGG 188
Db      2 TrpLeuIleuProIleuValIleuThrSerIleuAlaSerSerAlaThrTrp-----AlaGly 19
Qy      189 CTTTGTCTGAAGGCGGACAGAGAAACACGAGCTGGATGATTCAGAGGCAAGCAATC 248
Db      20 GlnProAlaSerProProValIalAepThrAlaGlnGlyAArgValIleuGlyIlyTrVal 39
Qy      249 ACTGTC-----CTGGGAAGCCCTGCTGCTGGAAGTGTCTCTGGAGTCCCTTGGCT 302
Db      40 SerIleuGluIleuValaIaPheThrGlnProValAlaValPheIleuGlyValProIleuAla 59
Qy      303 GCTCCCGGCTGGAGATCCCTGGCATTTTACGAACCCGACGCTGCATCGCCTGGAGTAAAC 362
Db      60 LysProProIleuGlySerIleuArgPheAlaProProGlnProAlaGluIleuProIleuSerPhe 79
Qy      363 TTGCGAAGAACCACTCTTACCTTAATTGTGCTTCAGAACTCA-----GAGTGGCTG 416
Db      80 ValIleuAsnThrThrSerTyrrProPrometCysCysGlnAspProValIalGluGlnMet 99
Qy      417 CTCTTAGAT-----CAACATCTCAAGTGCATTAACCGGAA 455
Db      100 ThrSerAspLeuPheThrAsnPheThrGlyLysGluArgLeuThrIleuIleuPhe----- 117
Qy      456 TTGAGAGTCAAGAGACTGCTCTTACGTAACATGTATGCGCCTGCCACCGGCAATACA 515
Db      118 -----SerGluAspCysIleuTyrgluAsnIleuTyThrProAlaAspLeuThrIly 134
Qy      516 GGCTCCAAAGCTCCCGCTGTGGTGTGTGTTCCAGAGAGTGCCTTCAAGATGGCTCAGGC 575
Db      135 ArgGlyArgLeuProValMetValTrpIleHsGlyGlyIleuValIleuGlyAla 154
Qy      576 TGCATTTTGAATGGGTCGCGCCCTGGCTGCTATGAGAAC-----GTGCTGTGTGGTTC 629
Db      155 PrometTyAspGlyValIalIleuAlaIalHsGluAsnPheThrValIalValAla 174
Qy      630 GTCCAGTACCGGCTAGGAATTTTGGTTCTTACACAGATGGATGAGATGCTCCGGG 689
Db      175 IleGlnTyArgIleuGlyIleTrpGlyPhePheSerThrGlyAspGluHsSerArgGly 194
Qy      690 AACTGGGCTTCAAGGACGAGGTGGCTGCTGTCTGCTGCTCCAGAGAACATCGAGTTC 749
Db      195 AsnTrpGlyHsIleuAspGlnValAlaIalHsIleuIleuIleuIleuIleuIleuIleu 214
Qy      750 TTGCGTGGGACCCGACGCTGTGACATCTTTGGCGATCC-----CGGGAGCCATA 803
Db      215 PheGlyGlyAspProGlySerValThrIlePheGlyGlyIleuSerPheThrIleGlyGlyIleu 234
Qy      804 AGTGTTCATGCTTATAGTGTCTCCAGGCAAGCTTATTCACAAAGCATCATG 863
Db      235 SerValSerValIleuValIleuSerProIleuAlaIlyAsnIleuPheThrIleuIleuSer 254
Qy      864 GAGAGTGGGTCGACATCCCTTACCTGAGAGCCCATGATTAAGAGAGAGAGAGAC 923
Db      255 GluSerGlyValAlaIleuThrValAlaIleuValaIlyAspMetIlyAlaIalIly 274
Qy      924 CTCGAGTGTTCGACATTTCTGTGTAAACAATGGCTGAGACTTGAGGCGCTGAGG 983
Db      275 GlnIleuAlaValIleuAlaGlyCysIlySerThrThrSerAlaIalPheThrPheValHs 294
Qy      984 TGCCTGAGGACAAACCTCCAGAGAGCTGTCGACCTCAGCGACGAAACAAAGTCTTTC 1043

```

[illegible]

```

STATE: Tennessee
COUNTRY: USA
ZIP: 37662-5075
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,515
FILING DATE: 09-SEP-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,879; 08/845,295
FILING DATE: 17-May-96; 25-April-97
ATTORNEY/AGENT INFORMATION:
NAME: Cheryl J. Tubach
REGISTRATION NUMBER: 38,346
REFERENCE/DOCKET NUMBER: 70432
TELECOMMUNICATION INFORMATION:
TELEPHONE: 423-229-6189
TELEFAX: 423-229-1239
INFORMATION FOR SEO ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: Amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-09-150-515-2

```

Alignment Scores:

Pred. No.:	1.02e-97	Length:	584
Score:	1091.50	Matches:	259
Percent Similarity:	54.39%	Conservative:	63
Best Local Similarity:	43.75%	Mismatches:	191
Query Match:	28.1%	Indels:	79
DB:	3	Gaps:	17

US-10-023-515-1 (1-2158) x US-09-150-515-2 (1-584)

QY	129	TGGGCTTTTTCCTGATGTTCCAGACCCCTGTGGGACACACAGCAGTGGGAAAACTGG	189
Db	2	TrpLeuLeuProLeuValLeuThrSerLeuAlaSerSerAlaThrTrp-----AlaGly	19
QY	189	CCTTCTCTAAGGGCCACAGAGAAACACAGCTGGGATGGATTCAAGGCAGCAAGTC	248
Db	20	GlnProAlaSerProProValValAlaSerThrAlaGlnGlyArgValAlaLeuGlyLysCty-Val	39
QY	249	ACTGTG-----CTGGAAAGCCCTGTGCCTGTGAACGTATTCCTCGAGAGTCCCTTGGT	302
Db	40	SerLeuLeuGlyLeuLeuAlaPheThrGlnProValAlaValAlaPheLeuGlyValProPheAla	59
QY	303	GCTCCGCCGCTGGGATCCCTGGCATTTACGAAACCCGAGCCGTGATCGCCTGGATAC	362
Db	60	LysProProLeuGlySerLeuAlaPheAlaProProGlnProAlaGlnProThrSerPhe	79
QY	363	TTGCGAAGAGCCACCTCCCTAATTTGGCTCCAGAACTCA-----GAGTGGCTG	416
Db	80	ValLysAsnThrThrSerLysProProMetCysGlnAspProValValGlnGlnMet	99
QY	417	CTCTTAAGT-----CACAACATGCTCAAGGTCATTGCCGGA	455
Db	100	ThrSerAspLeuPheThrAsnPheThrGlyLysGlnLysArgLeuThrLeuGlnPhe-----	117
QY	456	TTGCGAGTGTCCAGAGCTGCGCTTACCTTAACATCTATACGGCTGCCCCACGCCATACA	515
Db	118	-----SerGlnAspCysLeuLysLysLeuAsnLysLysThrProAlaAspLeuThrLys	134
QY	516	GGCTCCAAAGTCCCGCTTTGTGTGTGTTCCCAAGAGAGTCCCTTCAAGACTGAGTCAGCC	575
Db	135	ArgGlyAlaGlyLeuProValMetValTrpLLeuLysGlyGlyLysValLeuValLeuGlyGlyAla	154
QY	576	TGCATCTTTGATGGAGTCCGCTGAGCTGAGCTATAGAGAC-----CTGTGTGTTTGTGCT	620

```

Db      155  PrometTyrAspGluValValLeuAlaIahisGluAsnPheThrValValValAla 174
      630  GTCCAGTACCGGCTGGAATATTTGGTTTCTCACACATGGAGATGAGATCTCCGGG 689
Db      175  IlegIntYrArgLeuGlyLettPglYrPheSerThrGlyAspGluHiserArgGly 194
      690  AACTGGGCTTCAGACAGCAGGTGGCTGTCTGTGGGTTCAGAAAGAACATTCAGATT 749
      195  AenTrpGluHisLeuAspGluValAlaIahisLeuHiserThrValGluGluAsnLeuAla 214
      750  TTGGTGGGAAACCCAGCTCTGTACCATCTTTGGGAGTCC-----GGCGAGCCATA 803
      215  PheGlyGlyAspProGlySerValThrIlePheGlyGluSerPheThrAlaGlyGlyGlu 234
      804  AGTGTTCATGCTATATCTGTCTCCCAATGGCAAGCTTTTCCACAAACCATCATG 863
      235  SerValSerValLeuValLeuSerProLeuAlaLysAsnLeuPheHiserThrAlaIleSer 254
      864  GAGAGTGGGGTGGCCATCATCCCTTACCTGAGAGCCCATGATTAAGAGAGTGAAGAC 923
      255  GluSerGlyValAlaLeuThrValAlaLeuValArgLysAspMetLeuAlaIahisLys 274
      924  CTGCAAGTGTGTCACATTTCTGTGATCAATGGGTCAAGCTGAGGCGCTGCTGAGG 983
      275  GlnIleAlaValLeuAlaGlyCySerLysThrThrThrSerAlaValIlePheThrValHis 294
      984  TGCCTGAGAGCAAAACCTCCCAAGAGCTGAGACCTTCAGCCAGAAACAAAGCTTTC 1043
      295  CysLeuArgGlnLysSerGluAspGluLeuAspLeuThrLeuLysMetLysPheLeu 314
      1044  ACT-----CGAGTGGTGTAT 1058
      315  ThrLeuAspPheHisGlyAspGluAsnArgGluSerHiserProPheLeuProThrValAlaAsp 334
      1059  GGTGCTTCTTCTTAATGAGCTTAAGTCTAATGTTCTCAGAAA-----GCATTTAAA 1112
      335  GlyValLeuLeuProLysMetProGluGluIleLeuAlaGluLysAspPheThrPheAsn 354
      1113  GCATTTCTTCTCATATGAGAGTCAATAACAGAGTGGTCTGTGCTGCT----- 1166
      355  ThrValProLysIleValGlyLysAsnLysGlnLysPheGlyTrpLeuLeuProThrMet 374
      1167  -----ATGAGAGAGCTCCTGAGATCCCTGACT 1193
      375  MetGlyPheProLeuSerGluGlyLysLeuAspGlnLysThrAlaThrSerLeuLeuTrp 394
      1194  GAGCTCAACAGTCCCTTGCCTCATCTGATCAAAACATCTGCACATCCCGCTCAG 1253
      395  LysSer-----TyrProIleAlaAsnIle-----ProGlu 404
      1254  TATTTGCACCTTGTGGCTAATGATCTTCATGACAGAC-----TCC 1298
      405  GluLeuThrProValAlaIleThr-----PheThrAspLysTyrLeuGlyGlyThrAspAsp 422
      1299  CTGATCTGAATCCGAGACAGTCTTCTGAGCTTGTGAGATGATGTTCTTGTGGTCCCT 1358
      423  ProValLysLysLysAspLeuPheLeuAspMetGlyAspValAlaPheGlyValPro 442
      1359  GCATGATCACAGCTGATATACAGAGATGTGGTGCACCTGTCTATGATGTTT 1418
      443  SerValThrValAlaArgGlnHisArgAspAlaGlyAlaProThrTyrMetCysTrpLysPhe 462
      1419  CGGACCGGCTCAGTGTCTTGAAGAC-----ACGAAACCGGCTTTGTCAAGCCGAC 1472
      463  GlnTyrArgProSerPheSerSerAspLysPheThrLysProLysThrValIleGlyAsp 482
      1473  CACGCTGAATGAGTCCGCTTGTGTTCGGTGTGCTTCGAGAGGGAATTTGTTATG 1532
      483  HisGlyAspGluIlePheSerValPheGlyPheProLeuLeuLysGlyAsp----- 499
      1533  TTCGAAAGGACGACGAGAGAGAGATTATCAGACCGGAGATGATGAATATCTGGCT 1592

```

```

Db      500  -----AlaProGluGluGluValSerLeuSerLysThrValMetLysPheThrAla 516
      1593  ACCTTGTCTGAAACCGGGAATCTTAATGGAAACGACTGTCTCTGGGCA-----GCT 1646
      517  AenPheAlaArgSerGlyAsnProAsnGlyGluGlyLeuProHiserTrpPheThrMet 536
      1647  TATAATCTGACTGAGCACTTACCTCCAGCTGAGCTTGAACATGAGCTCCGACAGACTC 1706
      537  TyrAspGlnGluGluGlyTyrLeuGlnIleGlyValAenThrGlnAlaAlaLysArgLeu 556
      1707  AAGAAACCGGCTGAGATTTTGGACAGCACATC 1742
      557  LysGlyGluGluValAlaPheThrAsnAspLeuLeu 568
      568  -----
      569  -----
      570  -----
      571  -----
      572  -----
      573  -----
      574  -----
      575  -----
      576  -----
      577  -----
      578  -----
      579  -----
      580  -----
      581  -----
      582  -----
      583  -----
      584  -----
      585  -----
      586  -----
      587  -----
      588  -----
      589  -----
      590  -----
      591  -----
      592  -----
      593  -----
      594  -----
      595  -----
      596  -----
      597  -----
      598  -----
      599  -----
      600  -----
      601  -----
      602  -----
      603  -----
      604  -----
      605  -----
      606  -----
      607  -----
      608  -----
      609  -----
      610  -----
      611  -----
      612  -----
      613  -----
      614  -----
      615  -----
      616  -----
      617  -----
      618  -----
      619  -----
      620  -----
      621  -----
      622  -----
      623  -----
      624  -----
      625  -----
      626  -----
      627  -----
      628  -----
      629  -----
      630  -----
      631  -----
      632  -----
      633  -----
      634  -----
      635  -----
      636  -----
      637  -----
      638  -----
      639  -----
      640  -----
      641  -----
      642  -----
      643  -----
      644  -----
      645  -----
      646  -----
      647  -----
      648  -----
      649  -----
      650  -----
      651  -----
      652  -----
      653  -----
      654  -----
      655  -----
      656  -----
      657  -----
      658  -----
      659  -----
      660  -----
      661  -----
      662  -----
      663  -----
      664  -----
      665  -----
      666  -----
      667  -----
      668  -----
      669  -----
      670  -----
      671  -----
      672  -----
      673  -----
      674  -----
      675  -----
      676  -----
      677  -----
      678  -----
      679  -----
      680  -----
      681  -----
      682  -----
      683  -----
      684  -----
      685  -----
      686  -----
      687  -----
      688  -----
      689  -----
      690  -----
      691  -----
      692  -----
      693  -----
      694  -----
      695  -----
      696  -----
      697  -----
      698  -----
      699  -----
      700  -----
      701  -----
      702  -----
      703  -----
      704  -----
      705  -----
      706  -----
      707  -----
      708  -----
      709  -----
      710  -----
      711  -----
      712  -----
      713  -----
      714  -----
      715  -----
      716  -----
      717  -----
      718  -----
      719  -----
      720  -----
      721  -----
      722  -----
      723  -----
      724  -----
      725  -----
      726  -----
      727  -----
      728  -----
      729  -----
      730  -----
      731  -----
      732  -----
      733  -----
      734  -----
      735  -----
      736  -----
      737  -----
      738  -----
      739  -----
      740  -----
      741  -----
      742  -----
      743  -----
      744  -----
      745  -----
      746  -----
      747  -----
      748  -----
      749  -----
      750  -----
      751  -----
      752  -----
      753  -----
      754  -----
      755  -----
      756  -----
      757  -----
      758  -----
      759  -----
      760  -----
      761  -----
      762  -----
      763  -----
      764  -----
      765  -----
      766  -----
      767  -----
      768  -----
      769  -----
      770  -----
      771  -----
      772  -----
      773  -----
      774  -----
      775  -----
      776  -----
      777  -----
      778  -----
      779  -----
      780  -----
      781  -----
      782  -----
      783  -----
      784  -----
      785  -----
      786  -----
      787  -----
      788  -----
      789  -----
      790  -----
      791  -----
      792  -----
      793  -----
      794  -----
      795  -----
      796  -----
      797  -----
      798  -----
      799  -----
      800  -----
      801  -----
      802  -----
      803  -----
      804  -----
      805  -----
      806  -----
      807  -----
      808  -----
      809  -----
      810  -----
      811  -----
      812  -----
      813  -----
      814  -----
      815  -----
      816  -----
      817  -----
      818  -----
      819  -----
      820  -----
      821  -----
      822  -----
      823  -----
      824  -----
      825  -----
      826  -----
      827  -----
      828  -----
      829  -----
      830  -----
      831  -----
      832  -----
      833  -----
      834  -----
      835  -----
      836  -----
      837  -----
      838  -----
      839  -----
      840  -----
      841  -----
      842  -----
      843  -----
      844  -----
      845  -----
      846  -----
      847  -----
      848  -----
      849  -----
      850  -----
      851  -----
      852  -----
      853  -----
      854  -----
      855  -----
      856  -----
      857  -----
      858  -----
      859  -----
      860  -----
      861  -----
      862  -----
      863  -----
      864  -----
      865  -----
      866  -----
      867  -----
      868  -----
      869  -----
      870  -----
      871  -----
      872  -----
      873  -----
      874  -----
      875  -----
      876  -----
      877  -----
      878  -----
      879  -----
      880  -----
      881  -----
      882  -----
      883  -----
      884  -----
      885  -----
      886  -----
      887  -----
      888  -----
      889  -----
      890  -----
      891  -----
      892  -----
      893  -----
      894  -----
      895  -----
      896  -----
      897  -----
      898  -----
      899  -----
      900  -----
      901  -----
      902  -----
      903  -----
      904  -----
      905  -----
      906  -----
      907  -----
      908  -----
      909  -----
      910  -----
      911  -----
      912  -----
      913  -----
      914  -----
      915  -----
      916  -----
      917  -----
      918  -----
      919  -----
      920  -----
      921  -----
      922  -----
      923  -----
      924  -----
      925  -----
      926  -----
      927  -----
      928  -----
      929  -----
      930  -----
      931  -----
      932  -----
      933  -----
      934  -----
      935  -----
      936  -----
      937  -----
      938  -----
      939  -----
      940  -----
      941  -----
      942  -----
      943  -----
      944  -----
      945  -----
      946  -----
      947  -----
      948  -----
      949  -----
      950  -----
      951  -----
      952  -----
      953  -----
      954  -----
      955  -----
      956  -----
      957  -----
      958  -----
      959  -----
      960  -----
      961  -----
      962  -----
      963  -----
      964  -----
      965  -----
      966  -----
      967  -----
      968  -----
      969  -----
      970  -----
      971  -----
      972  -----
      973  -----
      974  -----
      975  -----
      976  -----
      977  -----
      978  -----
      979  -----
      980  -----
      981  -----
      982  -----
      983  -----
      984  -----
      985  -----
      986  -----
      987  -----
      988  -----
      989  -----
      990  -----
      991  -----
      992  -----
      993  -----
      994  -----
      995  -----
      996  -----
      997  -----
      998  -----
      999  -----
      1000  -----

```

[illegible]

```

QY      1572 AAGATGATGAATATCTGGGCTTACCTTTGCTCGAACCGGGAATCTTAAAGGAAAGCAAGCTG 1631
Db      468 MetValMetIysTrpAlaAsnPheAlaArgAsnGlyAsnProAsnGlyGluGlyIleu 487
QY      1632 TCTGTGGGACAGGTATTAATTCGACTGAGCGAGTACTCCAGCTGAGACTTGAACATGAGC 1691
Db      488 ProGlnTrpProAlaTrpAspTrpLysGluGlyTrpLeuGlnIleGlyAlaThrGln 507
QY      1692 CTCGAGACGAGACTCAAGAACCGCGGTGGATTTTGGACC 1733
Db      508 AlaAlaGlnIysLeuLysAspLysGluValAlaIlePheTrpThr 521

RESULT 13
US-08-446-100-26
; Sequence 26, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
; US-08-446-100-26

Alignment Scores:
Pred. No.: 3,446-83
Score: 942.50
Percent Similarity: 59.33%
Best Local Similarity: 44.94%
Query Match: 24.34%
DB: 3
Gaps: 6

QY      468 GAAGACTGCTTACCTGAACATCTATGCGCTCCGACCGACCATGACAGGCTCAAGCTC 527
Db      1 GlnHisCysLeuIysTrpAlaAsnIleGlyThrProAlaAspLeuThrIysValAsnArgClen 20

```

QY	528	CCCGCTGGAGTGGTGTCCSAGAGAGAGTCCCTTCAAGACCTGGATCCGCTCATCTGAT	587
Db	21	ProValMetValThrIleHisGlyIleGlyLeuMetValGlyAlaAlaSerThrTyrAsp	40
QY	588	GGATCCGCGCTGGTGGTCCATGAGACCTGGCTGGTGTGTGTCGTCAGTACCAGGCTAGAA	647
Db	41	GlyLeuAlaLeuAlaIleHisGlyLysMetValValValIleThrIleGlnTyrArgLeuIly	60
QY	648	ATATTGGTTTCTTCAACAATCGGATCAAGATGCTCCGGGAACTGGGACTTCAGAGAC	707
Db	61	IleTyrGlyPhePheSerThrGlyAspIleHisSerArgGlyAsnTyrGlyIleValAsp	80
QY	708	CAGGTGGCTGCTGTCTGGTCCAGAAAGAAATCGAGTTCTCCGGTGGGAGCCCAAGC	767
Db	81	GlnValAlaAlaLeuAlaGlyTyrValGlnAspAsnIleAlaSerPheGlyIleAsnProIly	100
QY	768	TCTGAGACATCTTTGGAGAGTCCGCGGAGACCATAGTCTTTTCACTTAACTGTCT	827
Db	101	SerValThrIlePheGlyIleSerAlaGlyIleGlySerValSerValLeuValLeuSer	120
QY	828	CCCATGGCCAAAGGCTTATTTCCAAACCCATCATGAGAGTGGGGTGGCCATCCCT	887
Db	121	ProLeuAlaLysAsnLeuPheHisArgAlaIleSerGlnSerGlyValAlaLeuThrSer	140
QY	888	TACCTGAGGCGCCATGATTATGAGAG-----ACTGAGACCTCGAGTGGTTCACAT	941
Db	141	ValLeuValIlyValGlyAspValIlyProLeuAlaGlnIleAlaIleThrAla---	159
QY	942	TTCTGTGGTAAACATGCGTCAGACCTCGAGGCGCTCGAGAGTGGCTGAGACAACACC	1001
Db	160	-----GlyCybIySerThrThrSerAlaAlaMetValHisCybLeuArgGlnIlyThr	177
QY	1002	TCCAAAGAGCTGCTGACCCCTCAGCCAGAA-----	1031
Db	178	GlnGlnGlnLeuLeuGlnThrThrLeuLysIleGlyAsnSerTyrLeuTyrThrTyrArg	197
QY	1032	-----ACAAAGCTTTCACTGAGAGTGTTGATGGTCTTTCTTCTAAT	1076
Db	198	GlnThrGlnAspGlnSerThrLeuLeuGlnIlyThrValIleAspIlyMetLeuLeuLys	217
QY	1077	GAGCCTCAGATCTATTGTCTCAGAAACATTTTAAAGCAATTCCTCATCATGCGAGTC	1136
Db	218	ThrProGlnGlnLeuGlnArgGlnTyrGlnAspPheHisIleThrValProTyrMetValGlyIle	237
QY	1137	AATTAACAAGAAGTGGCTCTCGTGGCTGCTGAGAGAGGCTCGCTGAGATCCTAGTGGC	1196
Db	238	AsnIlySerGlnIlyPheGlyTyrPheIleProMetGlnLeuMetSerTyrProLeuSerIle	257
QY	1197	TCC-----AACAAAGTCCCTTGGCTCCATCTGATCAAAAC-----ATCCTGAC	1241
Db	258	GlyGlnLeuAspGlnIlyThrAlaMetSerLeuGlnSerProIleProLeuPheAla	277
QY	1242	ATCCCGCTCAGTATTGGACCTTGTGGCTAATGAATCTTCATGACAAACACTCCCTG	1301
Db	278	IleAlaIySerGlnLeuIleIleProGlnAlaThrGlnIyIyTyrLeuGlyGlyThrAspAspThr	297
QY	1302	ACTGAATACGAGACAGTCTTCTGGACTTTGGTGGAGATGTGTTCTTGTGCTCCCTCA	1361
Db	298	ValIlyValIyAspLeuIleLeuAspPheIleAlaAspValMetPheGlyValProSer	317
QY	1362	CTGATCAAGCTGATATCACAAGATGCTGTGTGACCTGTCTACTTCTATGAGTTCCG	1421
Db	318	ValIleValAlaIleArgAsnHisArgAspAlaGlyIleProThrTyrMetTyrGlnPheGln	337
QY	1422	CACGGGCTCAGTGTGTTGAAGACACGAAGCGGCTTTGTCAAGCGGACACAGCTGAT	1481
Db	338	TyrArgProSerPheSerSerAspMetIyAspProIyThrValIleGlyAspHisGlyAsp	357
QY	1482	GAAATCCCTTGTGTGGTGTGTCCTTCTGAAAGGGGACATGTGTTATTTGGAAGA	1541
Db	358	GlnLeuPheSerValPheGlyAlaIleProPheLys-----GlnGly	371
QY	1542	GCCACGAGAGAGAAAGTAACTAGACCCGAAAGATGAATACTGGGTACTTTGCT	1601

Db	372	AlaSerGluGluGluGluIleAArgLeuSerIleMetValMetIlePheThrPheAlaAsnPheAla	391
QY	1602	CGAACCCGGGAATCTTAATGAGACGACCTGTCTCTGTGGCCAGCTTAATATGACTGAG	1661
Db	392	ArgLengIyaenProAengIlyLgIyLeuProhiStrpProGluYrAengInlygIlu	411
QY	1662	CAGNACTCCGCTGCGACTTGAAATAGATGACCTCGAGAGAGACCTCAAGAAACCCGGGTG	1721
Db	412	GlyIyrLeuGluInIleGlyAlaAlaAsnThrGlnAlaAlaGlnIlySeuIyBaapLygIuVal	431
QY	1722	GATTTTGACACGAC	1736
Db	432	AlaPheTrpIhrAsn	436
RESULT 14			
US-08-446-100-28			
; Sequence 28, Application US/08446100			
; Patent No. 6001625			
; GENERAL INFORMATION:			
; APPLICANT: Broomfield, Clarence A			
; APPLICANT: Willard, Charles B			
; APPLICANT: Lockridge, Okana			
; TITLE OF INVENTION: Site-directed Mutagenesis of Esterases			
; NUMBER OF SEQUENCES: 31			
; CORRESPONDENCE ADDRESSES:			
; ADDRESSEE: Hendricks and Assoc.			
; STREET: 9669 A Main Street, P.O. Box 2509			
; CITY: Fairfax			
; STATE: VA			
; COUNTRY: US			
; ZIP: 22031			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent in Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/446,100			
; FILING DATE: 19-MAY-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Hendricks, Glena			
; REGISTRATION NUMBER: 32,535			
; REFERENCE/DOCKET NUMBER: Broomfield			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (703) 425-4250			
; TELEFAX: (703) 425-2767			
; INFORMATION FOR SEQ ID NO: 28:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 454 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: unknown			
; MOLECULE TYPE: protein			
; HYPOTHEITICAL: YES			
; ANTI-SENSE: YES			
; FRAGMENT TYPE: N-terminal			
; ORIGINAL SOURCE:			
; ORGANISM: human esterases			
; US-08-446-100-28			
Alignment Scores:			
Pred. No.: 3,44e-83 Length: 454			
Score: 942.50 Matches: 200			
Percent Similarity: 59.55% Conservative: 65			
Best Local Similarity: 44.94% Mismatches: 149			
Query Match: 24.34% Indels: 31			
DB: 3 Gaps: 6			

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 08:53:20 ; Search time 148.694 Seconds

(without alignments)
11126.674 Million cell updates/sec

Title: US-10-023-515-1

Perfect score: 3872
Sequence: 1 ccaagcgcgcgaacacagc.....aaaaaaaaaaaaaaaaaa 2158

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 3420798

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.epool/6664091/runat_14062005_133501_9433/app.query.fasta_1.4238
-DB=published Applications AA -OFMT=faetan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOCL=0 -LOOEXT=0 -UNITs=bits -START=1 -END=-1 -MATRIX=bloum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=6664091@cgn2_1_1295@runat_14062005_133501_9433
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEOUTPUT -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	79.5	581	13 US-10-023-515-2	Sequence 2, Appli
2	3079	79.5	581	15 US-10-674-636-2	Sequence 2, Appli
3	3079	79.5	581	16 US-10-757-262-46	Sequence 46, Appli
4	2915	75.3	575	15 US-10-451-168-91	Sequence 91, Appli
5	2900	74.9	642	15 US-10-433-256-10	Sequence 10, Appli
6	2607.5	67.3	581	15 US-10-451-168-92	Sequence 92, Appli
7	2606	67.3	525	15 US-10-094-749-2375	Sequence 2375, Ap
8	2597.5	67.1	581	15 US-10-114-270-196	Sequence 196, App
9	2588.5	66.9	618	15 US-10-381-898-2	Sequence 2, Appli
10	2454	63.4	469	15 US-10-104-047-2219	Sequence 2219, Ap
11	1985	51.3	542	16 US-10-233-933A-2	Sequence 2, Appli
12	1985	51.3	542	16 US-10-233-933A-4	Sequence 4, Appli
13	1780	46.0	356	15 US-10-451-168-93	Sequence 93, Appli
14	1230	31.8	549	9 US-09-925-298-689	Sequence 689, App
15	1230	31.8	549	14 US-10-102-806-689	Sequence 689, App
16	1230	31.8	550	15 US-10-267-756-5	Sequence 5, Appli
17	1230	31.8	559	16 US-10-858-271-28	Sequence 28, Appli
18	1186.5	30.6	554	9 US-09-895-860-4	Sequence 4, Appli
19	1186.5	30.6	554	15 US-10-377-072-4	Sequence 4, Appli
20	1186.5	30.6	554	16 US-10-377-072-4	Sequence 4, Appli
21	1179.5	30.5	571	10 US-09-931-836-23	Sequence 23, Appli
22	1179.5	30.5	571	11 US-09-833-245-1090	Sequence 1090, Ap
23	1179.5	30.5	571	13 US-10-036-342-23	Sequence 23, Appli
24	1179.5	30.5	571	13 US-10-036-041-23	Sequence 23, Appli
25	1179.5	30.5	571	14 US-10-028-072-542	Sequence 542, App
26	1179.5	30.5	571	14 US-10-035-855-23	Sequence 23, Appli
27	1179.5	30.5	571	14 US-10-140-808-542	Sequence 542, App
28	1179.5	30.5	571	14 US-10-121-049-542	Sequence 542, App
29	1179.5	30.5	571	14 US-10-123-904-542	Sequence 542, App
30	1179.5	30.5	571	14 US-10-140-470-542	Sequence 542, App
31	1179.5	30.5	571	14 US-10-175-746-542	Sequence 542, App
32	1179.5	30.5	571	14 US-10-176-918-542	Sequence 542, App
33	1179.5	30.5	571	14 US-10-176-921-542	Sequence 542, App
34	1179.5	30.5	571	14 US-10-227-884-210	Sequence 210, App
35	1179.5	30.5	571	14 US-10-036-214-23	Sequence 23, Appli
36	1179.5	30.5	571	14 US-10-137-865-542	Sequence 542, App
37	1179.5	30.5	571	14 US-10-140-474-542	Sequence 542, App
38	1179.5	30.5	571	14 US-10-035-719-23	Sequence 23, Appli
39	1179.5	30.5	571	14 US-10-142-431-542	Sequence 542, App
40	1179.5	30.5	571	14 US-10-143-114-542	Sequence 542, App
41	1179.5	30.5	571	14 US-10-230-163-210	Sequence 210, App
42	1179.5	30.5	571	14 US-10-036-160-23	Sequence 23, Appli
43	1179.5	30.5	571	14 US-10-230-338-210	Sequence 210, App
44	1179.5	30.5	571	14 US-10-142-419-542	Sequence 542, App
45	1179.5	30.5	571	14 US-10-218-631-210	Sequence 210, App

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application us/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A. J.
; TITLE OF INVENTION: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2

Alignment Scores:
  Pred. No.:      3,396-269      Length:      581
  Score:          3079.00        Matches:      581
  Percent Similarity: 100.00%    Mismatches:  0
  Best Local Similarity: 100.00%  Gaps:        0
  Query Match:    79.52%        Indels:      0
  DB:             13            Gaps:        0

US-10-023-515-1 (1-2158) x US-10-023-515-2 (1-581)

QY  96  ATGCACAGGAGCTTACTGATCTGCTTCAACAAAGGCTTTTCTGATTCACAGCC 155
    |||
Db  1  MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleGlnPro 20

QY  156 CTGTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 215
    |||
Db  21  LeuLeuGlnIleArgGlnTrpGlyLysThrGlyProSerAlaGlnGlyProGlnAlaArg 40

QY  216 ACAAGGCTGGAGATTCAGAGGACAGCAAGTCACTGTGCTGGAAAGCCCTGTGCTGTG 275
    |||
Db  41  ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60

QY  276 AACGTGTTCTCGGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGGGATTTAGAAC 335
    |||
Db  61  AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80

QY  336 CCGGACCTGCATCGCCCTGGGATTAATTGGAGAAAGCACTCTCACTCAATTTTGTGC 395
    |||
Db  81  ProGlnProAlaSerProTrpAsnLeuArgGlnAlaThrSerTrpProAsnLeuGly 100

QY  396 CTCACAGACTCAGAGTGGCTGCTCTTAAGATCAACAATGCTCAAGGTGATTAACCGAAA 455
    |||
Db  101  LeuGlnAsnSerGlnTrpLeuLeuAsnArgGlnIleMetLeuLysValIleTrpProLys 120

QY  456 TTCGAGGTGCAGAGACTGCTTACCTGAACATCTATAGCGCTGCCACCGCATACA 515
    |||
Db  121  PheGlyValSerGlnAspGlySerLeuTrpLeuAsnIleTrpAlaProAlaIleAlaAspThr 140

QY  516 GGCTCAAGCTCCCGCTTGGTGTGTGCTCCAGAGAGTGCCTTTCAAGACTGCTCAAGCC 575
    |||
Db  141  GlySerLysLeuProValLeuValIleTrpPheProGlyGlyAlaPheLysThrGlySerAla 160

QY  576 TCCATCTTTTATGGGTCCGCTGCTGCTATGAGAGAGTGCCTGTTGTGGTGCATCAG 635
    |||
Db  161  SerIlePheAspGlySerAlaLeuAlaIleAlaTyrgLysAspValLeuValValValGln 180

QY  636 TACGGGCTAGAAATTTTGGTTTCTTCAACACATGGGATCAAGCATGCTCCGGGAACTGG 695
    |||
Db  181  TyrArgLeuGlyIlePheGlyPhePheThrThrTrpAspGlnIleAlaProGlyAsnTrp 200

QY  696 GCCTTCAGAGACAGAGTGGCTGCTGTCTGCGGTCCAGAAACATCGAATTTTGGCT 755
    |||
Db  201  AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnIleGlnPhePheGly 220

QY  756 GGGGACCCCACTCTGTGACCATCTTTGGCGAGTCCGCGGAGGCATTAAGTTTCTAGT 815
    |||
Db  221  GlysAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240

QY  816 CTTATACTGTCTCCATGGCCAAAGGCTTATTCACAAAGCATCAGAGAGTGGGGTG 875
    |||
Db  241  LeuIleLeuSerProMetAlaLysGlyLeuPheIleLysAlaIleMetGlnSerGlyVal 260

QY  876 GCCATCATCCCTTAACCTGAGAGCCCATGATTATGAGAAAGTGAAGACTGTCAGTGGT 935
    |||
Db  261  AlaIleIleProTrpLeuGlnAlaIleAlaAspTrpGlyLysSerGlnAspLeuGlnValVal 280

QY  936 GCACATTTCTGTGTACATAGCTCAGACTCTGAGGCCCTGCTGAGTCCCTGAGAGACA 995
    |||
Db  281  AlaIlePheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
```

```
QY  996 AAACCTCGAAGAGCTGTGACCTTACGCCGAGAAACAAAGTCTTTCACTGAGTGT 1055
    |||
Db  301  LysProSerLysGlnLeuLeuThrLeuSerGlnLysTrpLysSerPheThrArgValVal 320

QY  1056 GATGTGCTTTTCTTTCTTAATGAGCCTCAGATCATATTGCTCAGAAAGCATTTAAAGCA 1115
    |||
Db  321  AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340

QY  1116 ATTCCTTCATATGAGATGCAATTAACACAGAGTGTGCTTCTGCTGCTCATGAAGAG 1175
    |||
Db  341  IleProSerIleIleGlyValAsnAsnIleGlnCysGlyPheLeuLeuProMetLysGln 360

QY  1176 GCTCCTGAGATCTCAGTGGCTCCAAAGTCCCTTGCCCTTCATGATTAACAAACATC 1235
    |||
Db  361  AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuIleGlnAsnIle 380

QY  1236 CTGCACATCCCGCTCAGATTTTGGACCTTGTGCTAATGAAATTTCAATGACAAAGCAC 1295
    |||
Db  381  LeuIleIleProProGlnTrpLeuIleLeuValAlaAsnGlnTrpPheIleAspLysIle 400

QY  1296 TCCCTGACTGAATCCGACAGTCTTCTGACTTGTGGAATGTGTTCTTTGTGCTC 1355
    |||
Db  401  SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValIlePhePheValVal 420

QY  1356 CTTGCACTGATCAGACGCTCGATATACAGAGATGCTGTGACACTGTCTACTTATGAG 1415
    |||
Db  421  ProAlaLeuIleThrAlaArgTrpIleArgAspAlaGlyAlaProValTrpPheTrpGln 440

QY  1416 TTTCGGGACCGGCTCGAGTCTTTTGAAGACAGAGGCGGCTTTTGTCAAGCCGACAC 1475
    |||
Db  441  PheArgIleArgProGlnCysPheGlnAspThrLysProAlaPheValLysAlaAspIle 460

QY  1476 GCTGATGAAGTCCGCTTTTGTTCGATGCTGCTCCTGAAAGGGGAGCAATGTTATGTC 1535
    |||
Db  461  AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValIleMetPhe 480

QY  1536 GAAGGACCCACGAGAGGAGAAATTACTGAGCCGGAAGATGAATACTGGGCTACC 1595
    |||
Db  481  GlnGlyValAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetCysLysTrpAlaThr 500

QY  1596 TTGCTCGAAGCCGGAATCTTAATGGAAAGACCTGTCTGTGGCCACTTATAATCTG 1655
    |||
Db  501  PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520

QY  1656 ACTGAGCAGTACTCTCAAGCTGACATTGAACATGAGCCTTGAGACAGACTCAAGAACCG 1715
    |||
Db  521  ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540

QY  1716 CGGATGATTTTGGACAGACACATCCCGCTGATCTGTCTGCTCCGACATGCTCCAC 1775
    |||
Db  541  ArgValAspPheThrPheSerThrIleProLeuIleLeuSerAlaSerAspMetLeuIle 560

QY  1776 AGTCTCTTTTCTTCTTAATCTTCTCTCTCTCTCAAGCCTTTCTTTTCTTTGTGCT 1835
    |||
Db  561  SerProLeuSerSerLeuThrPheLeuSerLeuGlnInProPhePhePheCysAla 580

QY  1836 CCT 1838
    |||
Db  581  Pro 581

RESULT 2
US-10-674-636-2
; Sequence 2, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Siles-Santiago, Imaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYLSTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674, 636
; CURRENT FILING DATE: 2003-09-29
```

; PRIOR APPLICATION NUMBER: US/10/023, 515
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 60/256,369
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: 60/279,508
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-674-636-2

Alignment Scores:

Pred. No.:	3,396-269	Length:	581
Score:	3079.00	Matches:	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	79.52%	Indels:	0
DB:	15	Gaps:	0

US-10-023-515-1 (1-2158) x US-10-674-636-2 (1-581)

```

OY 96 ATGCCACAGGACTTACTCATCTGCTTCAACAATGGTGGCTTTTCTTGATTCACGCC 155
DB 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCyPhePheLeuLeuLeuGlnPro 20
OY 156 CTGTTGGGACACAGACAGTGGGGAATAACTGGGCTTCTGCTGAAGGCGCACAGAGAAC 215
DB 21 LeuLeuGlnYHLeuArgGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnArgAsn 40
OY 216 ACCAGGCTGGGATGGATTCAGAGGACAGCAAGTCACTGTGCTGGAGAGCCCTGCTGCTG 275
DB 41 ThrArgLeuGlyTrpTrpLeuGlnYHLeuGlnValThrValLeuGlySerProValProVal 60
OY 276 AACGTATCTCCGAGATCCCTTTGCTGCTCCCGCTGGAGATCCCTGAGATTACGAAC 335
DB 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
OY 336 CCGGAGCTGATGCGCCCTGGAGATACTTGAGAGAACCACTCTCAATATTGTC 395
DB 81 ProGlnProAlaSerProTrpAspPheLeuArgGlnAlaThrSerTrpProLeuLeuGlyS 100
OY 396 CTCACGAATCAGAGTGGCTGCTTCAATCAACATGCTCAAGGTGATTAACCGAAG 455
DB 101 LeuGlnAsnSerGlnTrpLeuLeuAspGlnHsMetLeuYHValHsIstYrProLys 120
OY 456 TTGGAGTGTCAAGAGACTGCTCTACCTGAACATCTATGCGCTGCGCACGCGCATACA 515
DB 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnHsIleTrpAlaProAlaHsIleAspHr 140
OY 516 GGCTCCAAAGCTCCCGCTTGTGTTGTTCCAGAGAGTGGCTTCAAGATGGCTCAAGCC 575
DB 141 GlySerTrpLeuProValLeuValTrpPheProGlyAlaPheLeuThrGlySerAla 160
OY 576 TCCATCTTTGAAGGGTCCGCTGCTGCTATGAGAGAGTGTGTGTGTCCTGCGAC 635
DB 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValValGln 180
OY 636 TACCGGCTAGGAATATTTGGTTCTTCAACACATGGAGATCAGATGCTCCGGGAACTGG 695
DB 181 TyrArgLeuGlyIlePheGlyPhePheTrpTrpAspGlnHsIleAlaProGlyAsnTrp 200
OY 696 GCCTTCAGAGACAGAGTGGCTGCTGTGCTGAGGTCAGAGAGACATCAGATTCTTCGGT 755
DB 201 AlapheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnHsIleGlnPhePheGly 220
OY 756 GGGGACCCCAAGCTCTGTGACATCTTTGGCAGAGTCCGGGAGCCCTAAGTGTCTTACT 815
DB 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
OY 816 CTTATACTGTCTCCATGCGCAAGGCTTATTCACAAAGCATCATGAGATGGGAGTG 875
  
```

```

DB 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHsIstYrAlaIleMetGlnSerGlyVal 260
OY 876 GCCATCATCCCTTACTGAGAGCCCATGATTTATGGAAGAGAGAGACCTGACGGTGT 935
DB 261 AlaIleIleProTrpLeuGlnAlaHsIleAspYrGlnLysSerGlnAspLeuGlnVal 280
OY 936 GCACATTTCTGTGTAAACAATGCGTCAGACTGTGAGAGCCCTGCTGAGTGGCTGAGACA 995
DB 281 AlaHsIlePheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuAspHr 300
OY 996 AAACCTCCAGAGAGCTGTGACCTTCAGCCAGAAACAAAGTCTTTCATCTGAGTGT 1055
DB 301 LysProSerTrpGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
OY 1056 GATGTCCTTTCTTTCTTAATGAGCTCTAAGATCTATGTCCTCAGAAACATTTAAAGA 1115
DB 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnLysAlaPheLysAla 340
OY 1116 ATTCCCTTCATCATGAGTCAATTAACAACAAGTGGGCTTCTGCTGCTGCTGAGAGAG 1175
DB 341 IleProSerIleIleGlyValAsnAsnHsIstYrCysGlyPheLeuLeuProMetLysGln 360
OY 1176 GCTCCTGAGATCCTGAGTGGCTCAACAAGTCCCTTGCCTCATCTGATTAACAAACATC 1235
DB 361 AlaProGlnIleLeuSerClySerAsnLysSerLeuAlaLeuHsIleuIleGlnAsnHsI 380
OY 1236 CTCACATATCCCGCTCAGATTTTGGACCTTGGCTTAAGATCTTCATGACAGAC 1295
DB 381 LeuHsIleProProGlnTrpLeuHsIleuValAlaAsnGlnTrpPheHsIleAspLysHs 400
OY 1296 TCCCTGACTGAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTTCTTTGGGTC 1355
DB 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheVal 420
OY 1356 CCTGACATGATCAAGACTGATATCAACAAGATGCTGTGACACTGTCTACTTATAGAG 1415
DB 421 ProAlaLeuIleIleThrAlaArgTrpHsIleAspAlaGlyAlaProValTrpPheTrpGln 440
OY 1416 TTTCGACACCGGCTCAGTGGCTTTGAAGACAGAGAGCCGGCTTTGTCAAACCGCACAC 1475
DB 441 PheArgHsIleArgProGlnCysPheGlnAspTrpLysProAlaPheValLysAlaAspHs 460
OY 1476 GCTGATGAAGTCCGCTTGTGTTGCTGGTGGTCTTCGGAAGGGGACATGTTATGTT 1535
DB 461 AlaAspGlnValAlaArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
OY 1536 GAAGGACCAACGAGAGAGAGAGATTACTGAGCCGGAAGATGATGAATATGGGCTAC 1595
DB 481 GlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpTrpAlaThr 500
OY 1596 TTTCCTCGAACCGGGAATCTTAATGGAACAACCTGTCTGTGGCCAGCTTAATATCTG 1655
DB 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
OY 1656 ACTGAGAGTACTCCTCAGCTGAGCTTGAACATGAGCTTCGACAGAGACTCAAGAACCG 1715
DB 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
OY 1716 CCGGTGATTTTGGACACAGACATCCCTGATCTGTCTGCTCGACATGCTCCAC 1775
DB 541 ArgValAspPheTrpPheThrIleProLeuIleLeuSerAlaSerAspMetLeuHs 560
OY 1776 AGTCTCTTTCTTCTTAACCTTCTCTCTCTCTGCTGAGCTTTCTTTTCTTTGTGCT 1835
DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
OY 1836 CCT 1838
DB 581 Pro 581
  
```

RESULT 3
 US-10-757-262-46

```
/ Sequence 46, Application US/10757262
/ Publication No. US20040197825A1
/ GENERAL INFORMATION:
/ APPLICANT: Karicheti, Venkateswarlu
/ APPLICANT: Eliabof, Scott D.
/ APPLICANT: Silos-Santlago, Imaculada
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
/ TITLE OF INVENTION: METABOLIC DISORDERS USING 44390, 54181, 211, 5687, 884,
/ TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
/ TITLE OF INVENTION: 33751, 56872, 14063, 20739, 32544, 43239, 44373, 51164,
/ TITLE OF INVENTION: 35010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
/ TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
/ TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
/ TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
/ TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
/ FILE REFERENCE: MP103-0071RNONAM
/ CURRENT APPLICATION NUMBER: US/10/757,262
/ PRIOR FILING DATE: 2004-01-14
/ PRIOR APPLICATION NUMBER: US 60/440,318
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US 60/444,783
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: US 60/457,901
/ PRIOR FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: US 60/468,775
/ PRIOR FILING DATE: 2003-05-08
/ PRIOR APPLICATION NUMBER: US 60/471,614
/ PRIOR FILING DATE: 2003-05-19
/ PRIOR APPLICATION NUMBER: US 60/478,742
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/488,529
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/491,156
/ PRIOR FILING DATE: 2003-07-30
/ PRIOR APPLICATION NUMBER: US 60/499,594
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 60/506,332
/ NUMBER OF SEQ ID NOS: 136
/ SOFTWARE: Faeseq for Windows Version 4.0
/ SEQ ID NO 46
/ LENGTH: 581
/ TYPE: PR1
/ ORGANISM: Homo sapiens
US-10-757-262-46
Alignment Scores:
Pred. No.: 3.39e-269 Length: 581
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.52% Indels: 0
Gaps: 0
DB: 16
US-10-023-515-1 (1-2158) x US-10-757-262-46 (1-581)
QY 96 ATGCACAGGAGCTACTTCATCTGCTTCACAAATGCTCTTTTCTGATTCACAGCC 155
DB 1 MetProGlnGlyLeuThrSerSerLaserGlnTrpCysPhePheLeuLLeuGlnPro 20
QY 156 CTGTGGGACACAGACAGTGGGGAATACTGGGCTTCTGCTGAAGGACACAGAGGAAC 215
DB 21 LeuLeuGlnLysAspGlnTrpGlyLysThrGlyProSerAlaGluGlyProGlnAspAsn 40
QY 216 ACCAGGCTGGGATGGATTCAGGCAAGCAAGTCACTGTGCTGGGAGCCCTGTGCTGTG 275
DB 41 ThrArgLeuGlyTrpLysGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
QY 276 AACGTGTTCTCGAGTCCCTTTGCTGCTCCCGCGTGGAGATCCCGGCAATTAAGAAC 335
DB 61 AsnValPheLeuGlyValPhePheAlaAlaPheProLeuGlySerLeuAlaGlnPheThrAsn 80
```

```
QY 336 CCGCAGCTTGATCCGCCCTGGGATTAATTGCGAAGACCCACCTCTACCTTAATTGTGC 395
DB 81 ProGlnProLaserProTrpAspLeuAlaGlnAlaThrSerLysProAsnLeuCys 100
QY 396 CTCACAGACTCAGAGGCTGCTCTAGATCAACACATGCTCAAGCTGATTAACCGAA 455
DB 101 LeuGlnAspSerGlnTrpLeuLeuLeuAspGlnLysMetLeuLysValLysTrpLys 120
QY 456 TTGGAGTGTCAAGAGATGCTGCTTACCTGAACATCTATGCGCTCCACGCGATACA 515
DB 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnLysLysAlaProAlaLysAlaAspThr 140
QY 516 GGTCTCAAGCTCCCGCTTGGTGTGTCCAGAGAGTGCCTTCAAGCTGCTCAAGCC 575
DB 141 GlySerLysLeuProValLeuValLysPheProGlyGlyAlaPheLysThrGlySerAla 160
QY 576 TCCATCTTTGATGGTCCGCCCTGGCTGCTATGAGACGTCGTGTTGTTGTTGTCAG 635
DB 161 SerLysPheAspGlySerAlaLeuAlaLysTrpGlnAspValLeuValValGln 180
QY 636 TACCGGCTAGGAAATTTGGTTTCTTACCAACATGGGATCAGATGCTCCGGGAACTGG 695
DB 181 TyrArgLeuGlyLysPheGlyPhePheThrTrpAspGlnLysAlaProGlyAsnTrp 200
QY 696 GCTTCAAGACACAGTGGCTGCTCTGCTCGGCTCAGAGAAACATCGAGTCTTCGCT 755
DB 201 AlaPheLysAspGlnValAlaAlaLeuSerTrpValGlnLysAsnLysGlnPheGly 220
QY 756 GGGGACCCGAGCTGTGACACATTTTGGAGAGTCCGCGGAGCCATAGTGTGTTAGT 815
DB 221 GlnAspProSerSerValThrLysPheGlyGlnSerLysAlaLysSerValLysSer 240
QY 816 CTATACCTGTCTCCATGACCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGCTG 875
DB 241 LeuLysLeuSerProMetAlaLysGlyLeuPheLysValAlaLysGlnSerGlyVal 260
QY 876 GCAATCATCCCTTACCTGAGAGGCCCATGATTATGAGAAAGTGAAGTGAAGTGTGTT 935
DB 261 AlaLysLysProLysLeuGlnLysAlaLysAspTrpGlnLysSerGlnAspLeuGlnVal 280
QY 936 GCAATTTCTGTGTGAACAATGAGTGAAGCTGAGAGCCCTGAGAGCTGCTGAGACA 995
DB 281 AlaLysPheCysGlyLysAsnAlaSerAspSerGlnAlaLeuLysValGlyValThr 300
QY 996 AAACCTTCAAGAGAGCTGCTGAACCTTCAAGCAAAACAAAGCTTTTCACTGAGTGT 1055
DB 301 LysProSerLysGlnLeuLeuThrLysSerGlnLysThrLysSerPheThrArgValVal 320
QY 1056 GATGTGCTTTCTTTCTTAATGAGCTCTAGATCTATGTTGTCAGAAAGCATTTAAACA 1115
DB 321 AspGlyAlaPhePheProAsnGlnTrpLeuLysPheLeuLysGlnLysAlaPheLysAla 340
QY 1116 ATTCTTTCATATGAGAGTCAATTAACACGAGTGTGCTTCTGCTGCTATGAGAG 1175
DB 341 IleProSerLysLysGlyValAsnAlaLysGlnLysGlyPheLeuLeuProMetLysGln 360
QY 1176 GCTCCGAGATCTAGAGGCTCTCAACAAAGTCCCTGCTCCATCTGATTAACAAATC 1235
DB 361 AlaProGlnLysLeuSerGlySerAsnLysSerLeuAlaLeuLysLysGlnAlaLys 380
QY 1236 CTGCACATCCGCGCTCAGATATTGACCTTGTGCTATGAAATCTTCAAGACAC 1295
DB 381 LeuLysLysLysProGlnTrpLysLysLysValAlaAsnGlnLysTrpLysAspLysVal 400
QY 1296 TCCCTGATGAAATCCAGAGACAGTCTTCTGACATTTGCTGAGATGTTGTTGTGTC 1355
DB 401 SerLeuTrpGlnLysLysAspSerLeuLysPheLysValLysPheValVal 420
QY 1356 CTTGACATGATCAGAGCTGATATCAAGAGATGCTGCTGCTGCTGCTGCTGCTGAG 1415
DB 421 ProAlaLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 440
QY 1416 TTTGCGACCGGCTCAGTGTCTTGAAGACAGAGCCGCTTTTGTCAAGCCGACAC 1475
```



```
|||||
Db 441 PheArgHisArgProGlnCysPheGluAerThrLysProAlaPheValLysAlaSerHis 460
Qy 1476 GCTGATGAAGTCCGCTTTGTGTTCCGTGCTGCTTCTGAAGGGGAGCAATGTATGTTTC 1535
Db 461 AAlaAPGluValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
Qy 1536 GAAGGAGCCAGGAGGAGGAGAAGTACTAGAGCCGGAAGATGATAAATACAGGCTAC 1595
Db 481 GUGUgAlaThrGluGluGluLysLeuLeuSerArgLysMetCysGlyTrrPalThr 500
Qy 1596 TTGCTCGAAGCCGGGAATCTTAATGGAAAGACTTCTCTGTGGCCAGCTTATAATCTG 1655
Db 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTyraLeu 520
Qy 1656 ACTGAGCACTACTCCAGCTGACTGAACATGAGCTTGAGACAGACTCAAGAACCG 1715
Db 521 ThrGluGlnTyrlenGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGluPro 540
Qy 1716 CGGGTGGATTTTGGACGACGACATCCGCTGATCTGTGCTGCGGACATGCTCCAC 1775
Db 541 ArgValaAPheThrPheThrPheSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
Qy 1776 AGTCCCTTTCTTCCCTTAATCTTCTCTCTCTCCAGCTTCTTTCTTTTGTGCT 1835
Db 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
Qy 1836 CCT 1838
Db 581 Pro 581

RESULT 4
US-10-451-168-91
; Sequence 91, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-451-168-91

Alignment Scores:
Pred. No.: 2,37e-254 Length: 575
Score: 2915.00 Matches: 553
Percent Similarity: 97.54% Conservative: 3
```

```
Beet Local Similarity: 97.02% Mismatches: 6
Query Match: 75.28% Indels: 8
DB: 15 Gaps: 1
US-10-023-515-1 (1-2158) x US-10-451-168-91 (1-575)
Qy 129 TGGTGGCTTTTCTGATATCTCCAGCCCTGTGGGACACAGACAGTGGGAAAACTGGG 188
Db 14 TrpAlaIleThrValLeuAlaAlaPro-----Thrylsgly 25
Qy 189 CTTTGTCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGCAAGTC 248
Db 26 ProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrrIleGlnIlysglnVal 45
Qy 249 ACTGTGCTGGGAAGCCCTGTGCTGTAACGTTCTCTCGAGTCCCTTGTGCTGCC 308
Db 46 ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAlaPro 65
Qy 309 CCGCTGGGATCCCTCGATTTACGAACCCGACGCTGCATCCGCTGGGATTAATTGCGA 368
Db 66 ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAsnLeuArg 85
Qy 369 GAAGCCACTCTTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTAATCAA 428
Db 86 GluAlaThrSerTyProAsnLeuCyLeuGlnAsnSerGluTrpLeuLeuAspGln 105
Qy 429 CACATGCTCAAGGTCAATTAACCGAAATTCGGAGTGCAGAACTGCTTACCTGAC 488
Db 106 HisMetLeuLysValHisTyProLysPheGlyValSerGluAspCysLeuTyrlenAsn 125
Qy 489 ATCTATGCGCTGCCGCCACGCGCATACAGGCTCCAGACTCCCGCTTGGTGTGCTTCCA 548
Db 126 IleTyrlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValTrpPhePro 145
Qy 549 GGAAGTGCCTTCAAGATGAGTCTCAGCTTCATCTTATGAGTCCGCTGCTGCTGCTAT 608
Db 146 GlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIlyr 165
Qy 609 GAGGACGTGCTGGTGTGCTGCTCCAGTACCGGCTAGAAATTTGGTTCCTTCCACNA 668
Db 166 GluAspValLeuValValValGlnTyArgLysGlyIlePheGlyPhePheThrThr 185
Qy 669 TGGGATCAGCATGCTCCGGGGAATGGGCTTCAAGSACAGGTGGCTGCTGCTGCTGG 728
Db 186 TrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrp 205
Qy 729 GTCCAGAAAGACATGAGTCTTCCGATGAGGACCCCACTGTGACCATCTTGGCGAG 788
Db 206 ValGlnLysAsnIleGluPhePheGlyAspProSerSerValThrIlePheGlyGln 225
Qy 789 TCCGCGGAGGACATAGTGTCTTACTTATCTGTCTCCATGGCCAAAGCTTATTC 848
Db 226 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerPrometAlaLysGlyLeuPhe 245
Qy 849 CACAAAGCATCATGAGAGTGGGAGGATCATCCCTTACCTGGAGGCCCATGATTAT 908
Db 246 HisLysAlaIleMetGlnSerIlyValAlaIleIleProTyrlenGlnAlaHisAspLys 265
Qy 909 GAGAAGAGTGAAGACCTGAGGTGTTGACATTTCTGTGTATACATGCGTCAAGCTCT 968
Db 266 GluLysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSer 285
Qy 969 GAGGCGCTGCTGAGGTGCTGAGGACAAACCTTCCAAAGAGCTGTGACCTTACGCCAG 1028
Db 286 GluAlaLeuLeuArgCysLeuArgThrLysProSerLysGluLeuLeuThrIleLeuSerGln 305
Qy 1029 AAAACAAAGCTTTTACCTCGAGTGTGATGATGCTTTCTTCTTAATGAGGCTTACAT 1088
Db 306 LysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 325
Qy 1089 CTATTGTCTCAGAAAGATTAAAGCAATTCCTTCATCATCGAGTCAATTAACGAGAG 1148
Db 326 LeuLeuSerGlnLysAlaIlePheLysAlaIleProSerIleIleGlyValAsnAsnHisGln 345
```

QY 1149 TGTGGCTTCTGCTGCTATGAGAGAGCTCTGAGATCTGATGCTTCAAGCTCC 1208
DB 346 CysGlyPheLeuLeuProMetLeuSegLysAlaProGluLysLeuSerGlySerAsnLysSer 365
QY 1209 CTGGCCCTCATCTGATACAAACATCTGACATCCGCTCAGTATTTGACCTGTG 1268
DB 366 LeuAlaLeuHisLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 385
QY 1269 GCTAAATGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
DB 386 AlaLeuGluLysPheHisAspLysHisSerLeuThrGluLysAspSerLeuLeuAsp 405
QY 1329 TTGCTTGAAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1388
DB 406 LeuLeuGlyAspValPhePheValValProAlaLeuLysLeuLysLeuLysLeuLys 425
QY 1389 GCTGGTGAACCTGCTTACTTCTTATGAGATTTGCGACCGGCTCAGTCTTGAAGACG 1448
DB 426 AlaGlyAlaProValLysPheLysGluPheArgHisArgProGlnCysPheGluAspThr 445
QY 1449 AAGCCGGCTTTTGTCAAGCCGACGCTGATGATGATGATGATGATGATGATGATGAT 1508
DB 446 LysProAlaPheValLysValAspHisAlaAspGluValArgPheValPheGlyGlyAla 465
QY 1509 TTCCGTAAGGGGACATTTGTTATGTTGTAAGAGACCGACGAGAGAGAGAGAGAGAG 1568
DB 466 PheLeuLysGlyAspLysLeuValMetPheGluGlyAlaThrGluGluGluLysLeuLys 485
QY 1569 CGAAGATGATGAAATAGTGGGCTTACTGCTGATGATGATGATGATGATGATGATGAT 1628
DB 486 ArgLysMetMetLysTrpTrpAlaThrPheAlaArgHisGlyAsnProAsnGlyAsnAsp 505
QY 1629 CTGCTCTGCGGCGACGTTTAAATCTGACTGACGACGATCCTCCAGCTGATGATGATG 1688
DB 506 LeuSerLeuTrpProAlaTrpAsnLeuThrGluGluLysLeuLysLeuLysLeuLys 525
QY 1689 AGCCTGAGACAGACGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1748
DB 526 SerLeuGlyGlnArgLeuLysGluProArgValAspPheTrpHisSerThrLeuProLeu 545
QY 1749 ATCTGCTGCTGCTGACATGCTCCAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1808
DB 546 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 565
QY 1809 CTCGAGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1838
DB 566 LeuGlnProPhePhePheCysAlaPro 575

RESULT 5
US-10-433-256-10
Sequence 10, Application US/10433256
Publication No. US20040081960A1
GENERAL INFORMATION:
APPLICANT: SANJANMALA, Madhusudan M.; YAO, Monique G.
APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
APPLICANT: ARVIZU, Chandra S.; RING, HuiJun Z.
APPLICANT: LEE, Ernestine A.; DING, Li
APPLICANT: HAFALIA, April J.A.; TANG, Y. Tom
APPLICANT: YU, Henry; TRIBOLEV, Catherine M.
APPLICANT: LU, Dyrung; Aina M.; LAL, Preeti G.
APPLICANT: WARREN, Bridget A.; YANG, Junning
APPLICANT: CHAMLA, Nandinder K.; NGUYEN, Daniel B.
APPLICANT: GANDHI, Ameena R.; LU, Yan
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-0313 USN
CURRENT APPLICATION NUMBER: US/10/433,256
CURRENT FILING DATE: 2003-05-30
PRIOR APPLICATION NUMBER: PCT/US01/47429
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: US 60/254,308
PRIOR FILING DATE: 2000-12-08

QY 1149 TGTGGCTTCTGCTGCTATGAGAGAGCTCTGAGATCTGATGCTTCAAGCTCC 1208
DB 346 CysGlyPheLeuLeuProMetLeuSegLysAlaProGluLysLeuSerGlySerAsnLysSer 365
QY 1209 CTGGCCCTCATCTGATACAAACATCTGACATCCGCTCAGTATTTGACCTGTG 1268
DB 366 LeuAlaLeuHisLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 385
QY 1269 GCTAAATGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
DB 386 AlaLeuGluLysPheHisAspLysHisSerLeuThrGluLysAspSerLeuLeuAsp 405
QY 1329 TTGCTTGAAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1388
DB 406 LeuLeuGlyAspValPhePheValValProAlaLeuLysLeuLysLeuLysLeuLys 425
QY 1389 GCTGGTGAACCTGCTTACTTCTTATGAGATTTGCGACCGGCTCAGTCTTGAAGACG 1448
DB 426 AlaGlyAlaProValLysPheLysGluPheArgHisArgProGlnCysPheGluAspThr 445
QY 1449 AAGCCGGCTTTTGTCAAGCCGACGCTGATGATGATGATGATGATGATGATGATGAT 1508
DB 446 LysProAlaPheValLysValAspHisAlaAspGluValArgPheValPheGlyGlyAla 465
QY 1509 TTCCGTAAGGGGACATTTGTTATGTTGTAAGAGACCGACGAGAGAGAGAGAGAGAG 1568
DB 466 PheLeuLysGlyAspLysLeuValMetPheGluGlyAlaThrGluGluGluLysLeuLys 485
QY 1569 CGAAGATGATGAAATAGTGGGCTTACTGCTGATGATGATGATGATGATGATGATGAT 1628
DB 486 ArgLysMetMetLysTrpTrpAlaThrPheAlaArgHisGlyAsnProAsnGlyAsnAsp 505
QY 1629 CTGCTCTGCGGCGACGTTTAAATCTGACTGACGACGATCCTCCAGCTGATGATGATG 1688
DB 506 LeuSerLeuTrpProAlaTrpAsnLeuThrGluGluLysLeuLysLeuLysLeuLys 525
QY 1689 AGCCTGAGACAGACGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1748
DB 526 SerLeuGlyGlnArgLeuLysGluProArgValAspPheTrpHisSerThrLeuProLeu 545
QY 1749 ATCTGCTGCTGCTGACATGCTCCAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1808
DB 546 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 565
QY 1809 CTCGAGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1838
DB 566 LeuGlnProPhePhePheCysAlaPro 575

Alignment Scores:
Pred. No.: 5,69e-253 Length: 642
Score: 2900.00 Matches: 548
Percent Similarity: 99.82% Conservative: 2
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 74.90% Indels: 0
DB: 15 Gaps: 0

US-10-023-515-1 (1-2158) x US-10-433-256-10 (1-642)

QY 186 GGGCTTCTGCTGAGAGGCGACAGAGAGACAGAGCTGGATGATGATGATGATGATGATGAT 245
DB 92 GlyProSerAlaGluGluGluProGlnArgAsnThrArgLeuLysTrpLysGln 111
QY 246 GTCACTGCTGGAGAACCTCTGTGCTGTGAGAGTGTCTCGAGATCCCTTGTGCTGT 305
DB 112 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAla 131
QY 306 CCCCCTGGAGATCCCTGCGATTTAGAACCCGACGCTGATGATGATGATGATGATGATGAT 365
DB 132 ProProLeuGlySerLeuArgPheTrpAsnProGlnProAlaSerProTrpAsnLeu 151
QY 366 CGAAGACCACTCTCAATTTGCTGCTGAGAGTGTCTCGAGATCCCTTGTGCTGT 425
DB 152 ArgGluAlaThrSerTrpProAsnLeuGlyLeuGlnAsnSerGluTrpLeuLeuAsp 171
QY 426 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGACTGCTTACCTG 485
DB 172 GlnHisMetLeuLysValHisTrpProLysPheGlyValSerGluAspCysLeuTrpLeu 191
QY 486 AACATCTATGCGCTGCGACCGACGATACAGGCTCCAGAGCTCCGCTGCTGCTGCT 545
DB 192 AsnLysTrpAlaProAlaHisAlaAspHisGlySerLysLeuProValLeuValTrpPhe 211
QY 546 CCAAGAGGCTTCTCAAGCTGCTGACGCTTCATCTTGTATGAGTCCGCTGCTGCTGCT 605
DB 212 ProGlyGlyAlaPheLysThrGlySerAlaSerLysPheAspGlySerAlaLeuAla 231
QY 606 TATGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
DB 232 TyrGluAspValLeuValValValValGlnTrpArgLeuGlyLysPheGlyPheThr 251
QY 666 AATGAGATCAAGCATGCTCCGAGAACTGGGCTTCAAGACAGAGTGGCTGCTGCTGCT 725
DB 252 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSer 271
QY 726 TGGCTCGAAGAACATGAGATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 785
DB 272 TrpValGlnLysAsnLysLeuLysPheGlyGlyAspProSerValTrpLysPheGly 291
QY 786 GAGTCCGCGGAGACATGAGTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
DB 292 GluSerAlaGlyAlaLysSerValSerSerLeuLeuSerProMetAlaLysGlyLeu 311

846 TTCCAAAGCCATGAGAGAGGAGGAGCCATCTTACCTGAGAGCCCATGAT 905
Db Phehtslvslallllemecsluserglvalalallelelprotylneuglulahlbaap 331
Qy 906 TATGGAAGAGTGAAGACCTGACAGGTTGGACATTTCTGTGTGAACAATCGCTGAC 965
Db 332 TGTGTluysserGlulabpneuglvalalahlani:spheCysglYAsnbnslaserabp 351
Qy 966 TCTGAGGCGCTGAGGAGGCTGAGGAGCAAAACCTCCAAAGAGCTGTGACCTGACG 1025
Db 352 SerGlulalaleuLeuargyLeuakrglhybProserLybgluLeuLeuthrLeuser 371
Qy 1026 CAGAAACAAAGCTTTTCACTGAGTGTGTGATGAGTGTCTTTCTTAATAGGCTCTTA 1085
Db 372 GlnystrlrysserPhehtlrglvalalabpGlYalabPheProabnslurProlen 391
Qy 1086 GATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATATGAGAGTCAATTAACAC 1145
Db 392 AspleuLeuSerGlulabPheLybvalalleProserllelleglyalabnbnhl: 411
Qy 1146 GAGTGTGCTTCCTGCTGCTTAAGAGAGGCTCCGAGATCTGAGTGTGCTCCAAACAG 1205
Db 412 GlnCybglYpPheLeuLeuProheLysglulalProglulvalLeuSerclYserbnlyb 431
Qy 1206 TCCCTTGGCCCTCATCTGATACAAACATCTGCAATCCGCGCTCAGTATTGGACCTT 1265
Db 432 SerleuAlaleuHlaleuHlleglnbnlileuHlslleProProglInlyrleuHlbleu 451
Qy 1266 GTGGCTAATGAATCTTCATGACAGACATCCCTGACTGAATTCGAGACAGTCTTCTG 1325
Db 452 ValalabnGluyr:Phehtl:abpLybhl:serLeuHnGlulabrgsberLeuLeu 471
Qy 1326 GACTTGTCTGAGAGTGTCTTGTGTGCTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTG 1385
Db 472 AspleuLeuGlYasPvalabPhevalalProalaleuHl:ethalabrgylxhl:baag 491
Qy 1386 GATGCTGTGACACCTGTACTTCTATGAGTTCGAGACCGGCTCAGTCTTTGAAGAC 1445
Db 492 AsplabglYalabProvalalYr:PhehtYrGlulabPheahghl:baagProglInCyberhebluabp 511
Qy 1446 ACGAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCTGCTG 1505
Db 512 ThrlybProalabPhevalalYbvalabPhehl:baabrglvalabrgPhevalabPheglYglY 531
Qy 1506 GCTTCTCGAAGGGGACATTTGTTATGTTGCAAGAGCCAGGAGAGAGAGAGTACTG 1565
Db 532 AlabPheLeuLybglYasPlalevalmecPheglulYalathrglulglulYbaleuLeu 551
Qy 1566 AGCCGGAAGATGATAAATTAAGTGGGCTACCTTGTGCTGAAACGGGAAATCTTAATGGGAAC 1625
Db 552 SerahglYbMetMeluLybYr:TrpalathrPhehl:ah:ghnGlYAsnProabnslYbAsn 571
Qy 1626 GACCTGTCTGTGTGCGACAGCTTAATTAATCTGACTGAGACAGTACTCAGCTGAGCTTGAAC 1685
Db 572 AspleuPheLeuTrpProalalYr:AsnLeuHnrglInlyrLeuGlInleuabpLeuabn 591
Qy 1686 ATGAGCTTCGAGACAGACATCAAGAACCGGGGAGTATTTGAGACAGACACATCCCC 1745
Db 592 MetSerleuGlYbInargYleuLybglulProargvalabgluPheTrpTh:serThrllePro 611
Qy 1746 CTGATCTGTGTGCTCGACAGTCTGCAAGTCTCTTCTTAATCTTCTGCTCT 1805
Db 612 LeuHlLeuSerAlabSerabPheMetLeuHl:serProLeuSerSerLeuThrPheLeuSer 631
Qy 1806 CTCCTCAGGCTTTTCTTTTCTTTTGTGCTCT 1838
Db 632 LeuLeuGlInProPhePhePheCysalabPro 642

RESULT 6
US-10-451-168-92
; Sequence 92, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:

APPLICANT: SMITHKLINE BEECHAM CORPORATION
APPLICANT: SMITHKLINE BEECHAM P.L.C.
APPLICANT: GLAXO GROUP LIMITED
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP50039
CURRENT APPLICATION NUMBER: US/10/451,168
CURRENT FILING DATE: 2003-11-12
PRIOR APPLICATION NUMBER: PCT/US01/49232
PRIOR FILING DATE: 2000-12-17
PRIOR APPLICATION NUMBER: 60/256,710
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/257,048
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 60/260,482
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: 60/264,922
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/266,797
PRIOR FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 60/276,988
PRIOR FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 60/281,535
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/289,622
PRIOR FILING DATE: 2002-06-28
NUMBER OF SEQ ID NOS: 110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 92
LENGTH: 581
TYPE: PRT
ORGANISM: Homo sapiens
US-10-451-168-92
Alignment Scores:
Pred. No.: 1.64e-226 Length: 581
Score: 2607.50 Matches: 499
Percent Similarity: 92.67% Conservative: 7
Best Local Similarity: 91.39% Mismatches: 25
Query Match: 67.34% Indels: 15
Gaps: 3
US-10-023-515-1 (1-2158) x US-10-451-168-92 (1-581)
Qy 129 TGTGCTTTTCTGATCTTCAGCCCTGTTGGACACAGACAGTGGGAAAACTGGG 188
Db 14 TrpalalaltrPvalabPhealabPro-----ThrlYsgly 25
Qy 189 CTTTGTGTAAGGGCCACAGAGAACACAGGCTGGAGTTCAGGGCAAGCAAGTC 248
Db 26 ProserAlabglulYrProglInabrgbnThrlrgleuGlYrPlleGlInglYbglInval 45
Qy 249 ACTGTGTGGAAAGCCCTGTGCTGTGAAGCTGTCTCGAGAGTCCCTTGTGCTCTCC 308
Db 46 ThrlValleuGlYserProvalabProvalabnvalabPheleuGlYvalabProPhehlalabPro 65
Qy 309 CCGGTGGATCCCTGCGATTTACGAACCGGAGCTGATGCGCTGGAGTAACTTGGGA 368
Db 66 ProleuGlYserleuabrgPhehtlrglYbAsnProglInProalabserProTrPabPbnleuabrg 85
Qy 369 GAAGCCACCTCTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTCTTAGATCA 428
Db 86 GlulalathrserYrProabnleuCybLeuGlInbnbnserGlultrPleuLeuabnrgln 105
Qy 429 CACATGCTGAAGTGCATTAACCCGAAATTCGAGTGTCAAGAAAGTGCCTCTTACTGAAC 488
Db 106 Hl:metLeuLybvalabInl:etYr:ProLybPheglYalabSerGlulabP:CysleuLytrleuabn 125
Qy 489 ATCTATGGGCTGCGCCACGCGGATTAACAGGCTCAAGGCTCCCGCTTGTGTGCTTCCCA 548
Db 126 Hl:etYrAlabProalabnslabPbnThrlYserLybLeuProvalabLeuvalaltrPbnPro 145
Qy 549 GGAGGTGCTTCAAGACTGAGCTCAGCTCTCATCTTTGATGGAGTCCGCTGCTGCTAT 608

QY 369 GAAGCACTCTCAACCTAATTGTCCTCAAGAACTCAGAGTGGCTGCTTAAATCA 428
 DB 86 GUAaathSerTyPProasnhLeuCybLeuGlnasSerGltUTrpleuLeuLeuaspGln 105
 QY 429 CACATGCTCAAGGTGATTAACCGAAATTCCGAGTGCAGAAAGATGCTCTTACCTGAAC 488
 DB 106 HismetLeuLyValHisTyPProLybPheGlyValIsSerGluabPybLeuTyTrleuasn 125
 QY 489 ATCTATGCCCTCCGCCCAACGGCTCAAGAGTCCCGCTCTTGCTGTGCTTCCCA 548
 DB 126 ILeTyAlaProAlaHisAlaasPThrGlySerLybLeuProValLeuValITrPhePro 145
 QY 549 GGAGGTGCTTCAAGACTGGCTCAGGCTCCATCTTGTATGGGTCGCGCCCTGGCTGCTAT 608
 DB 146 GLyGlyAlaPheLeuTyTrIleSerAlaSerITlePheaspGlySerAlaLeuAlaAlaTyP 165
 QY 609 GAGGACGTGCTGTGTGTGTCGTCCAGTAACCGCTAGAAATTTGGTTTCTTCAACACA 668
 DB 166 GUAaPValLeuValValValGlnTyPArgLeuGlyITlePheGlyPhePheThrTr 185
 QY 669 TGGATTCAGCAATGCTCCGGGAACTGGGCTTCAAGAACCAAGGTGGCTGCTGTCTGG 728
 DB 186 TTPaSPGlnHisAlaProGlyAsnTrpAlaPheLybaspGlnValAlaAlaLeuSerTyP 205
 QY 729 GTCCAGAAAGAAATGAGTCTTCCGTGGGGACCCGAGCTGTGTGACATCTTTGGCGAG 788
 DB 206 ValGlnLybAsnITleGluPhePheGlyGlyaspProSerSerValIThrITlePheGlyGln 225
 QY 789 TCCGGCGGAGCCATAAGTCTTCTAGTCTTATAGTCTTCCATGCGCAAGGCTTATTC 848
 DB 226 SerAlaGlyAlaITleSerValSerIleuITleLeuSerProMetAlaLybGlyLeuPhe 245
 QY 849 CACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACTGGAGGCCCATGATTAT 908
 DB 246 HisLybAlaITleMetGlySerGlyValAlaITleITleProTyTrleuGlnAlaHisaspTyP 265
 QY 909 GAGGAAGAGTGAAGACCTCGAGGTGGTGGACATTTCTGTGGTAACAATCGTCAAGCTCT 968
 DB 266 GlnLybSerGlnaspLeuGlnValAlaAlaHisPheCybGlybAsnAsnAlaSeraspSer 285
 QY 969 GAGGCCCTGCTGAGGTGCTGAGAACAAACCTCCAGAGAGCTGTGAACCTCAGCCAG 1028
 DB 286 GlnAlaITleLeuArgCybLeuArgIThrLybProSerLybGlnLeuLeuIThrLeuSerGln 305
 QY 1029 AAAACAAAGTCTTTCATCTGAGTGGTGAATGGTCTTTCTTCTTATAGAGCTCTAGAT 1088
 DB 306 LybIThrLybSerPheIThrArgValValaaspGlyAlaPhePheProaspGluProLeuasp 325
 QY 1089 CTAATGCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACCAACGAG 1148
 DB 326 LeuLeuSerGlnLybAlaPheLybAlaITleProSerITleITleGlyValaAsnAsnHisGln 345
 QY 1149 TGTGGCTTCTGCTGCTGCTGAAGAGGCTCCTGAGATCCTGAGTGGCTCAACAAGTCC 1208
 DB 346 CybGlyPheLeuLeuProMetLybGlnAlaProGlnITleLeuSerGlySerIThrLybSer 365
 QY 1209 CTTGGCCCTTCATCTGATACAACAAACATCTGCAATCCGCGCTCAAGTATTTGGACCTTGTG 1268
 DB 366 LeuAlaITleHisITleGlnAsnITleHisITleProProGlnIThrITleHisITleVal 385
 QY 1269 GGTATGAAATATCTTCATGACAGACATCCCTGACATCGAAATCCGAGACAGTCTTGGAGC 1328
 DB 386 AlaaspGlnTyPTrPheITHisaspLybHisSerIThrGlnITleArgaspSerIThrLeuasp 405
 QY 1329 TTGCTTGAGATGTGTCTTGTGTGCTCCTGACATGATCAGAGCTGATATCAGAGAT 1388
 DB 406 LeuLeuGlyaspValPhePheValValProAlaITleITleThAlaArgIThrHisArg--- 424
 QY 1389 GCTGTGACACCTGTCTACTTCTATGAGTTTGGACACCGGCTCAAGTCTTTGAAGACAGC 1448
 DB 424 ----- 424

QY 1449 AAGCCGGCTTTTGTCAAAACCCAGCACGCTGATGAATGCCGTTGTGTTCCGTGGTGCC 1508
 DB 424 ----- 424
 QY 1509 TTCTCGAAGGGGACATTGTTATGTTTCAGAGAGCCACGAGAGAGAAATTACTGAGC 1568
 DB 425 -----GluGlyAlaThrGluGluGluLybLeuLeuSer 435
 QY 1569 CGAAGATGATGAATPACTGGGCTACCTTGTGTCGACCGGGAAATCCTAATGGGAACGAC 1628
 DB 436 ArgLybMetMetLybTyTrITrPAlaThrPheAlaArgIThrGlybAsnProaspGlybAsnasp 455
 QY 1629 CTGTCTGTGCGCCAGCTTATATCTGACTGAGAGTCACTCACTGAGCTGATGAAACATG 1688
 DB 456 LeuSerIThrITrProAlaIThrAsnLeuIThrGlnIThrLeuGlnLeuaspLeuaspMet 475
 QY 1689 AGCTTCGACAGACACTCAAGAACCCGGGTGATTTTGGACACGACCATCCCTCTG 1748
 DB 476 SerLeuGlyGlnArgLeuLybGluProArgValaaspPheITrPheIThrITleProLeu 495
 QY 1749 ATCTGTCTGCTCCGACATGCTTCACAGTCTCTTCTTCTTACTTCTCTCTCTC 1808
 DB 496 ITleLeuSerAlaSeraspMetLeuHisSerProLeuSerSerIThrIThrPheLeuSerIThr 515
 QY 1809 CTCACGCTTCTTCTTCTTCTTGTGCTCTCT 1838
 DB 516 LeuGlnIThrPhePhePheCysAlaPro 525
 RESULT 8
 US-10-114-270-196
 ; Sequence 196, Application US/10114270
 ; Publication No. US20040030110A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Malvankar, Uriel M.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Patcurajan, Meera
 ; APPLICANT: Liu, Zhaozhong
 ; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Li, Li
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Zelnusen, Bryan D.
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Shenooy, Suresh G.
 ; APPLICANT: Pena, Carol E.A.
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Burgess, Catherine E.
 ; APPLICANT: Gerlach, Valerie
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: Gangolli, Bsha A.
 ; APPLICANT: Taupier Jr., Raymond J.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Lietze, Mario W.
 ; APPLICANT: Raastelli, Luca
 ; APPLICANT: Edinger, Shlomo R.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: MacDougall, John R.
 ; APPLICANT: Rothenberg, Mark B.
 ; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-322C
 ; CURRENT APPLICATION NUMBER: US/10/114,270
 ; PRIOR FILING DATE: 2002-11-27
 ; PRIOR APPLICATION NUMBER: 60/281,086
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: 60/281,136
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: 60/281,863
 ; PRIOR FILING DATE: 2001-04-05

/	PRIOR APPLICATION NUMBER:	60/281,906
/	PRIOR FILING DATE:	2001-04-05
/	PRIOR APPLICATION NUMBER:	60/282,020
/	PRIOR FILING DATE:	2001-04-06
/	PRIOR APPLICATION NUMBER:	60/282,930
/	PRIOR FILING DATE:	2001-04-10
/	PRIOR APPLICATION NUMBER:	60/282,934
/	PRIOR FILING DATE:	2001-04-10
/	PRIOR APPLICATION NUMBER:	60/283,512
/	PRIOR FILING DATE:	2001-04-12
/	PRIOR APPLICATION NUMBER:	60/283,710
/	PRIOR FILING DATE:	2001-04-13
/	PRIOR APPLICATION NUMBER:	60/284,234
/	PRIOR FILING DATE:	2001-04-17
/	Remaining Prior Application data removed - See File Wrapper or PALM.	
/	NUMBER OF SEQ ID NOS:	470
/	SEQ ID NO:	196
/	LENGTH:	581
/	TYPE:	PRT
/	ORGANISM:	Homo sapiens
/	US-10-114-270-196	
Alignment Scores:		
Pred. No.:	1,32e-225	Length:
Scores:	2597.50	Matches:
Percent Similarity:	95.07%	Conservative:
Best Local Similarity:	94.31%	Mismatches:
Query Match:	67.08%	Indels:
DB:	15	Gaps:
US-10-023-515-1 (1-2158) x US-10-114-270-196 (1-581)		3
QY	186 GGGGCTTCTGTGAAGGCCACAGAGAACAACAGCGTGATGATTCAAGGCACAA	245
Db	27 GlyProSerAlaGluGlyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyValGln	46
QY	246 GTACCTGTGCGGGAGACCCTGCGCCCTGTAACGTGTTCTCCGAGTCCTCCCTTTGCTGCT	305
Db	47 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValPropheAlaIa	66
QY	306 CCCCCGTGGGATCCCTGCAGATTACGAACCCGACGCTCATCGCCCTGGATACTTG	365
Db	67 ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpAspAsnLeu	86
QY	366 CGAAGAACCACTCCCTACCCCTAATTGTCGCTCAGAACTCAGAGTGGCTGCTTAGAT	425
Db	87 ArgGlnAlaThrSerTyProAsnLeuCybLeuGlnAsnSerLutRpeLeuLeuAsp	106
QY	426 CAACACATCTCAAGGTGATTACCCGAATTGAGAGTGCAGAGAATGCGCTTACCTG	485
Db	107 GlnHMetLeuIyValHisTyRProLybPheGlyValSerGlnAspCybLeuTyzLeu	126
QY	486 AACATATAGCCGCTGCCAAGCCGATTAACAGGCTCCAAGCTCCCGTCTTGATGCTTC	545
Db	127 AsnIleTyAlaProAlaHisAlaAspTrpGlySerTybLeuProValLeuAlaTrpPhe	146
QY	546 CCAGAGGTGCTTCAGACTGAGCTCAGGCTCATCTTTGATGGGCGCCCTGGCTGCG	605
Db	147 ProGlyGlyAlaPheTyThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa	166
QY	606 TATAGAGACGCTGCTGTTGTCCTCCAGTACCGGCTTAGAGAAATTTGGTTCTTCAAC	665
Db	167 TyrGlnAspValLeuValValValGlnTyArgLeuGlyIlePheGlyPhePheThr	186
QY	666 ACATGGATCAGCATGCTCCGGGGAACATGGGCTTCAGAGCACAGAGTGGCTGCTGCC	725
Db	187 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheTybAspGlnValAlaAlaLeuSer	206
QY	726 TGCGTCCAGAGAACATGAGTTCTTCCGATGGGGAACCCCAAGCTCTGTGACCATTTGGC	785
Db	207 TrpValGlnTybAsnIleGlyPhePheGlyGlyAspProSerSerValThrIlePheGly	226
QY	786 GAGTCGCGGAGCCATGATGTTTCTAGCTTATACGTCTCCCAAGGCCAAAGGCTTA	845

[illegible]

APPLICANT: BOROMSKY, Mark L.; DING, Li;
APPLICANT: DUGGAN, Brendan; ELIOT, Vicki S.;
APPLICANT: GANDHI, Ameena R.; GRIFPIN, Jennifer A.;
APPLICANT: HARALIA, April J.A.; ISON, Craig H.;
APPLICANT: KHAN, Farrah A.; LAU, Preeti G.;
APPLICANT: LEE, Ernestine A.; LO, Dung Anna M.;
APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
APPLICANT: SANTANALA, Madhusudan M.;
APPLICANT: TANG, Y. Tom; TRIBOUDET, Catherine M.;
APPLICANT: CHANLA, Narinder K.; WALSH, Roderick T.;
APPLICANT: WARREN, Bridget; XU, Yuming;
APPLICANT: YANG, Junming; YAO, Monique; YUE, Henry
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-023 USN
CURRENT APPLICATION NUMBER: US/10/381, 898
CURRENT FILING DATE: 2003-10-17
PRIOR APPLICATION NUMBER: PCT/US01/30662
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/236, 947
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: US 60/238, 864
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/242, 323
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/247, 581
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/249, 519
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/252, 834
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/250, 567
PRIOR FILING DATE: 2000-11-30
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7473645CD1
US-10-381-898-2
Alignment Scores:
Pred. No.: 8,88e-225 Length: 618
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 66.85% Indels: 23
DB: 15 Gaps: 1
US-10-023-515-1 (1-2158) x US-10-381-898-2 (1-618)
QY 186 GGGCTTGGCTGAGAGGGCCAGAGAAACACAGGCTGGATGAGTTAGAGGCAAGCAA 245
DB 92 GTPProseRAlaGluGlyPProGlnAArgAenThrAArgLeuGlyTrrpIleGlnGlybGln 111
QY 246 GTCACTGTGCTGGGAAGCCCTGTGCTGTGAACGTGTTCTCGAGTCCCTTGTGCTGCT 305
DB 112 ValThrValLeuGlySerProValProValAenValPheLeuGlyValProPheAlaAla 131
QY 306 CCCCCGCTGGGATCCCTGGGATTTACGAACCCGAGCCTGTGATCGCCCTGGGATTAATTG 365
DB 132 ProPheLeuGlySerLeuAArgPheThrAenProGlnProAlaSerProTrrAAspAenLeu 151
QY 366 CGAGAGCCACCTCCATCACTAATTGCTGCTCGAAGTCAAGAGTGGCTGCTTAAGT 425
DB 152 ArgGluAlaThrSerTrrProAenLeuCybLeuGlnAenSerGluTrrPheLeuLeuAAsp 171
QY 426 CAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAGACTGCTTAACCTG 485
DB 172 GlnHleMetLeuLeuValHisTrrProLybPheGlyValSerGluAAspCybLeuTrrLeu 191

QY 486 AACATCTATGCGCTGCCACGCGCATACAGGCTCCAAAGTCCCGCTTGGTGTGCTTC 545
DB 192 AenIleTrrAlaProAlaHisAlaAspThrGlySerLybLeuProValLeuValTrrPhe 211
QY 546 CCGAGAGGTGCTTCAAGACTGGCTCAAGCTTCATCTTGTATGAGTCCGCGCTGCTGCC 605
DB 212 ProGlyGlyAlaPheLybThrGlySerAlaSerIlePheAAspGlySerAlaLeuAlaAla 231
QY 606 TATGAGACGTGCTGTGTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 665
DB 232 TrrGluAAspValLeuValValValGlnTrrAArgLeuGlyIlePheGlyPhePheThr 251
QY 666 ACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGAGCAAGAGTGGCTGCTGCTCC 725
DB 252 ThrTrrAAspGlnHisAlaProGlyAenTrrAlaPheLybAAspGlnValAlaAlaLeuSer 271
QY 726 TGGGTCCAGAAAGACATGAGTTCTTGGGTGGGACCCCAAGCTGTGTGACCATCTTGGC 785
DB 272 TrrValGlnLybAenIleGluPhePheGlyGlyAAspProSerSerValThrIlePheGly 291
QY 786 GAGTCCGGGGGAGCCATAGTGTCTTATAGTCTTATAGTCTTCCATGGCCAAAGCTTA 845
DB 292 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaLybGlyLeu 311
QY 846 TTCACAAGCCATCATGAGAGAGTGGGCTCATCTCCCTTACCTGGAGGCCCATGAT 905
DB 312 PheHisLybAlaIleMetGluSerGlyValAlaIleIleProTrrLeuGlnAlaHisAsp 331
QY 906 TATGAGAAAGTGAAGACCTCGAGGTGTGGACATTTCTGTGTATCAATGCTGCAGAC 965
DB 332 TrrGluLybSerGluAAspLeuGlnValValAlaHisPheCybGlyAAspAAspAAsp 351
QY 966 TCTGAGGCTTCTGAGGTGCTGAGGACAAACCTTCAAGAGACTGTGACCTCAGC 1025
DB 352 SerGluAlaLeuLeuAArgCybLeuAArgThrLybProSerLybGluLeuLeuThrLeuSer 371
QY 1026 CAGAAACAAAGTCTTTCATCTGAGTGTGATGAGGTCTTCTTCTTAATGAGCTTCA 1085
DB 372 GlnLybThrLybSerPheThrAArgValAlaAAspGlyAlaAAspPheProAAspProLeu 391
QY 1086 GATCTATTGTCTCAGAAAGCATTTTAAGCAATTCCTTCATCATGCGAGTCAATTAACAC 1145
DB 392 AAspLeuLeuSerGlnLybAlaPheLybAlaIleProSerIleIleGlyValAAspAAspHis 411
QY 1146 GAGTGTGCTTCTGCTGCTGCTTATGAAGAGCTCCTGAGATCTCAGTGGCTCCAACAG 1205
DB 412 GluCybGlyPheLeuLeuProMet----- 419
QY 1206 TCCCTTGGCCCTCCATCTGATACAAAAACATCCGACATCCGCGCTCAGATTTGACCTT 1265
DB 420 -----HisIleProGlnTrrLybLeuHis 428
QY 1266 GTGGCTAATGATATCTTCATGACAGCACTCCCTGACTGAATCCAGACAGTCTTGTG 1325
DB 429 ValAlaAAspGlnTrrPheHisAspLybHisSerLeuThrGlnIleAArgAspSerLeuLeu 448
QY 1326 GACTTGTGAGATGTGTTCTTGTGTGCTGCTGCACTGATCAAGCTGCTGATTAACAGA 1385
DB 449 AAspLeuLeuGlyAAspValPhePheValAlaProAlaLeuIleThrAlaArgTrrHisArg 468
QY 1386 GATGTGTGAGCAGCTGTACTCTTATAGATTTGGGACCGGCTCAGAGCTTGTGAAGC 1445
DB 469 AAspAlaGlyAlaAAspValTrrPheTrrGluPheAArgHisAArgProGlnCybPheGlyAAsp 488
QY 1446 ACAGAGCGGCTTGTGTCAAGCGCAAGCTGATGAAGTCCGCTTGTGTGTGCTGCTGCT 1505
DB 489 ThrLybProAlaPheValLybAlaAspHisAlaAAspGlnValAArgPheValPheGlyGly 508
QY 1506 GCTTCTCGAAGGGGAGCACTTGTATGTTCCAGAGGACCAAGAGAGAGAGAACTTACG 1565
DB 509 AlaPheLeuLybGlyAAspIleValMetPheGlnGlyAlaThrGlnGlnGlyLybLeuLeu 528

FILE REFERENCE: SHIG FP02US006
CURRENT APPLICATION NUMBER: US/10/233, 933A
CURRENT FILING DATE: 2002-09-03
PRIOR APPLICATION NUMBER: JP2002-057908
PRIOR FILING DATE: 2002-04-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 542
TYPE: PRT
ORGANISM: Felis catus
US-10-233-933A-2

Alignment Scores:
Pred. No.: 3,62e-170 Length: 542
Score: 1985.00 Matches: 373
Percent Similarity: 83.21% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 51.27% Indels: 2
Gaps: 2

US-10-023-515-1 (1-2158) x US-10-233-933A-2 (1-542)

```
QY 186 GGGCCCTTGTCTGTAAGGGCCACAGAGAAACACAGGCTGGATGATTCAAGGCAAGCA 245
Db 22 GlyProAlaAlaAspAlaProValArgSerThrArgLeuGlyTyrValArgGlyLeuGln 41
QY 246 GTCACTGTCTGGGAAGCCCTGTGCTGTGAACGTGTTCTTGAGAGTCCCTTCTGCT 305
Db 42 ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyIleProTyrAlaAla 61
QY 306 CCCCCGCTGGGATCCCTGCGATTATACGAACCCGACGCTGCATGCGCCGGAATACTTG 365
Db 62 ProProLeuGlyProLeuArgPheLeuGlnProValProAlaLeuProGlyAsnAspPhe 81
QY 366 GGAAGACCACTCTTACCTTAATTGTGCTTCCAGAACTCAAGATGCTGCTCTTAGAT 425
Db 82 ArgAsnAlaThrSerTyrProLeuLeuCySPheGlnAspLeuGlnTyrPheValSerTyr 101
QY 426 CAACATGCTCAAGGTCAATTAACCGAAATTCGAGAGTGTCAAGAACTGCTCTTACTG 485
Db 102 GlnHisValLeuLeuValArgTyrProLeuGlnAlaSerGlnAspCysLeuTyrLeu 121
QY 486 AACATCATATGCGCGCCACGCGGATACAGGCTCCAGGCTCCCGCTGTGCTGCTGCT 545
Db 122 AsnIleTyrAlaProAlaHisAlaAspAsnGlySerAsnLeuProValMetValTyrPhe 141
QY 546 CCAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTGAATGGTCCGCTGCTGCTGCC 605
Db 142 ProGlyGlyAlaPheLeuMetGlySerIleSerSerPheAspGlySerAlaLeuAlaAla 161
QY 606 TATGAGACGTGCTGTGTTGTGTGTCAGTACCGGCTTGAATATTTGGTTTCTTACC 665
Db 162 TyrGlnAspValLeuIleValIleThrThrGlnTyrArgLeuGlyIlePheGlyPheAsp 181
QY 666 ACATGGATATGATGCTCCGCGGGAACGCGGCTTCAAGACCGAGTGGCTGCTGCTCC 725
Db 182 ThrGlnAspGlnHisAlaArgGlyAsnTyrAlaLeuLeuAspGlnValAlaAlaLeuThr 201
QY 726 TGGGTCCAGAGAACTGAGTCTTCCGTTGGGGAACCCAGCTCTGTGACCATTTTGGC 785
Db 202 TyrValArgAspAsnIleGlnPhePheGlyGlyAspProAspSerValThrIlePheGly 221
QY 786 GAGTCCCGGGAGCCATTAAGTGTCTTATAGTCTTATCTCTCCATGCGCAAGGCTTAA 845
Db 222 GluSerIleGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAsnGlyLeu 241
QY 846 TTCCACAAAGCATCATGAGAGTGGGCTGACATCATCCCTTACCTG---GAGGCCAT 902
Db 242 PheHisValAlaIleMetGluSerGlyValAlaIleLeuProLeuLeuMetArgProPro 261
QY 903 GATTATGAGAGAGTGAAGCTGACGAGTGTGTCACATTTCTGTGTAACATTCCTCA 962
Db 903 GATTATGAGAGAGTGAAGCTGACGAGTGTGTCACATTTCTGTGTAACATTCCTCA 962
```

```
Db 262 GlyAspGlnArgGlyLeuAspLeuGlnValLeuAlaArgIleCysGlyCysHisAlaSer 281
QY 963 GACTCTGAGGCCCTGCTGAGGCTGCTGAGCAAAACCTCCAGAGACTGCTGACCTTC 1022
Db 282 AspSerAlaAlaLeuLeuGlnCysLeuArgAlaLysProSerGlnGluLeuMetAspIle 301
QY 1023 AGCCAGAAAACAAAGCTTTTCACTCGAGTGGTGAATGGTGTCTTTCTTAATGAGCCT 1082
Db 302 SerTyrLeuLeuThrPheSerIleProValIleAspAspPhePheProAspIlePro 321
QY 1083 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCATTCCTTCCATCATGAGAGTCAATAC 1142
Db 322 ValAlaLeuLeuThrGlnHisAlaPheAsnSerValProSerIleIleGlyValAsn 341
QY 1143 CACGAGTGTGCTTCTGCTGCTGCTGCTGATGAAGAGGCTGCTGAGTCTGAGGCTCCAC 1202
Db 342 HisGlnCysAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlyCysSerAsn 360
QY 1203 AAGTCCCTTGGCTTCATCTGATACAAACATCTCCATCCGCTCAGATTTTGGAC 1262
Db 361 ArgSerLeuAlaLeuTyrLeuValHisThrPheLeuAsnIleProThrGlnTyrLeuHis 380
QY 1263 CTGTGCTTAATGAATTAATCTTCATGACACAGCATCTCTGACTGAATAATCCAGACAGTCTT 1322
Db 381 LeuValAlaAspHisTyrPheTyrAsnLysHisSerProValGlnIleArgAspSerPhe 400
QY 1323 CTGACATTCCTTGAAGATGTTCTTGTGCTGCTGCTGCTGACATGATACAGCTCGATATAC 1382
Db 401 LeuAspLeuLeuGlyAspValLeuPheValProGlyValIleThrAlaArgTyrHis 420
QY 1383 AGAGATGCTGTGACCTGCTCTACTTATGATGTTTCGACCGGCTCAAGTCTTTGAA 1442
Db 421 ArgAspAlaGlyAlaProValTyrPheTyrGlnPheGlnHisProProGlnCysLeuAsn 440
QY 1443 GACACGAAGCCGCTTTTGTCAAAAGCCGACCACTGATGAAGTCCGCTTGTGTGCT 1502
Db 441 AspThrArgProAlaPheValIleAspHisSerAspGlnIleArgPheValPheGly 460
QY 1503 GTGCTCTTCCGAGGCGGGAATTTGTTTCCAAAGAGCCAGGAGAGAGAAAGTTA 1562
Db 461 GlyAlaPheLeuLysGlyAspIleValMetPheGlnGlyAlaTnGlnGlnGlyLeu 480
QY 1563 CTGAGCCGGAAGATGAATTAATCTGAGCTACTTGTGCTGACCGGAACTTATGCGG 1622
Db 481 LeuSerArgLysMetMetArgTyrTyrAlaAsnPheAlaAspGlnGlyAspProAsnGly 500
QY 1623 AACGACTGTCTGTGTGCGCAGCTTATATCTGACTGACAGAGTACCTTCAGCTGACTTG 1682
Db 501 GluGlyValProLeuTyrProAlaTyrThrGlnSerGlnGlnTyrLeuLysLeuAspLeu 520
QY 1683 AACATGAGCTTCGACAGAGACTCAAGAGACCGCGGCTGATTTTGGACAGCACCATC 1742
Db 521 SerValSerValGlyGlnLysLeuLysGlnGlnGlnValGlnPheTyrMetAsnThrIle 540
```

RESULT 12
US-10-233-933A-4
Sequence 4, Application US/10233933A
Publication No. US2004021417A1
GENERAL INFORMATION:
APPLICANT: Yamazaki, Masao
TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
FILE REFERENCE: SHIG FP02US006
CURRENT APPLICATION NUMBER: US/10/233, 933A
PRIOR APPLICATION NUMBER: JP2002-057908
PRIOR FILING DATE: 2002-04-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 542
TYPE: PRT
ORGANISM: Felis catus

US-10-233-933A-4
Alignment Scores:
Pred. No.: 3,62e-170 Length: 542
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 51.27% Indels: 2
DB: 16 Gaps: 2
US-10-023-515-1 (1-2158) x US-10-233-933A-4 (1-542)
QY 186 GGGCTTCTGTGAAGGCGCAAGAGAACACAGAGCTGTGAGTGGATTCAGGCAAGCA 245
DB 22 GTPYAAIAAIAAPAIAProvalAArgserThnrgleuGlyTrpValAArgGlybGln 41
QY 246 GTCACTGTGGGAAGCCCTGTGCTGTGAACGTGTCTCTGGAGTCCCTTGTGCT 305
DB 42 ThnThrValLeuGlySerThnThrValProValAAsnMetPheLeuGlyIleProTyAlaAla 61
QY 306 CCCCCTGGGATCCCTGCATTTACGAACCCGAGCTGCATCGCCCTGGATTAACCTG 365
DB 62 ProProLeuGlyProLeuAArgPheLeuGlnProLyProAlaLeuProGlyAAsnAAspHe 81
QY 366 CGAGAACCCACCTCTTACCTTAATTTGTGCTTCCAGAACTCAGAGTGGCTCTTAGAT 425
DB 82 ArgAsnAlaThrSerTyrrProLybLeuCybPheGlnAAspLeuGlnTrpLeuValSerTyrr 101
QY 426 CAACACATGCTCAAGGTGATTAACCCGAATTCGAGGTGCAGAGAGTCCCTCTTACTG 485
DB 102 GlnHleValLeuLybValAArgTyrrProLybLeuGlnAlaSerGlnAAspCybLeuTyrrLeu 121
QY 486 AACATCTATGCGCTGCGCCACGCGCATACAGGCTCCAGAGTCCCGCTTGTGTGTTC 545
DB 122 AsnAlleTyrrAlaProAlaHleAlaAAspAAsnGlySerAAsnLeuProValMetValTrpPhe 141
QY 546 CCAAGAGTGCCTTCAAGACTGGCTCAGCTCATCTTGAAGGTCCGCTGCTGCTGCC 605
DB 142 ProGlyGlyAlaPheLeuMetGlySerAlaSerSerPheAAspGlySerAlaLeuAlaAla 161
QY 606 TATAGACAGT 665
DB 162 TyrrGlnAAspValLeuIleValThnThrGlnTyrrArgLeuGlyIlePheGlyPheAAsp 181
QY 666 ACATGGGATCAGAGTCCCGGGGAACCTGGCTTCAAGAACAGAGTGGCTGTGTCC 725
DB 182 ThnGlyAAspGlnHleAlaArgGlyAAsnTrpAlaLeuAAspGlnValAlaAlaLeuThr 201
QY 726 TGGGTCCAGAGAACATCGAGTTCTTGGTGGGAGCCGAGCTGTGTGACATCTTGGC 785
DB 202 TrpValAArgAAspAAsnIleGlnPhePheGlyGlyAAspProAArgSerValThnIlePheGly 221
QY 786 GAGTCCGGGGAGCCATAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 845
DB 222 GlnSerAlaArgGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAAsnGlyLeu 241
QY 846 TTCCCAAAAGCCATCATGAGAGTGGGTGGGAGCCATCCCTTACCTG---GAGGCCAT 902
DB 242 PheHleTyrrAlaIleMetGlnSerGlyValAlaIleLeuProLeuMetAArgProPro 261
QY 903 GATTATGAAGAAGTGAAGACCTCGAGGTGTGTGACATTTCTGTGTAAACAATGCTCA 962
DB 262 GlnAAspGlnAArgLybLybAAspLeuGlnValIleuAlaArgIleCybGlyCybHleAlaSer 281
QY 963 GACTCTGAGCCCTGTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
DB 282 AspSerAlaAlaLeuLeuGlnCybLeuAArgAlaLybSerProSerGlnGlnLeuMetAAspIle 301
QY 1023 AGCCAGAAAACAAAGCTTTTCACTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1082
DB 302 SerTyrrLybLeuThnThrPheSerIleProValIleAAspAAspPhePheProAAspGlyPro 321
QY 1083 CTAGATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCCATCATCGAGTCAATTAAC 1142

DB 322 ValAlaLeuLeuThnThrGlnLybAlaPheAAsnSerValProSerIleIleGlyValAAsnAAsn 341
QY 1143 CACGAGTGTGGCTTCTGTGCTGCCTTGAAGAGGCTCTGAGATCTCGAGTGCCTCAAC 1202
DB 342 HleGlnCybAlaPheLeuLeu---SerThnGlnPheSerGlnIleLeuGlyGlySerAAsn 360
QY 1203 AAGTCCCTGGCCCTGCATGTGTATCAAAAACATCCCGACATCCCGCCCTCAGTATTTGCA 1262
DB 361 ArgSerLeuAlaLeuTyrrLeuValIleThnThrPheLeuAAsnIleProThnGlnTyrrLeuHle 380
QY 1263 CTGTGGCTTAATGAATCTTCTCATGACAGACACTCCCTGACTGAATCCGAGACAGTCTT 1322
DB 381 LeuValAlaAAspHleTyrrPheTyrrAAsnLybHleSerProValGlnIleAAspSerPhe 400
QY 1323 CTGAGACTGTGAGAGT 1382
DB 401 LeuAAspLeuLeuGlyAAspValLeuPheValAlaProGlyValAlaThnAlaArgTyrrHle 420
QY 1383 AGAGATGTGTGACACTGTCTACTTATGAGTTTGGCAGCCGCGCTCAGTGTGTGA 1442
DB 421 ArgAAspAlaGlyAlaProValTyrrPheTyrrGlnPheGlnHleSerProGlnCybLeuAAsn 440
QY 1443 GACAGGAAGCCGCTTTTGTGTCAAGCCGACACGCTGATGAAGTCCGCTTGTGTGTGTGT 1502
DB 441 AAspThnAArgProAlaPheValLybAlaAAspHleSerAAspGlnIleAArgPheValPheGly 460
QY 1503 GTGTCTTCTGAAAGGGGACATGTGTATGTGTGAAGAGACCAAGGAGGAGAGATTA 1562
DB 461 GlnAlaPheLeuLybGlyAAspIleValMetPheGlnGlyAlaThnGlnGlnGlnLybLeu 480
QY 1563 CTGAGCCGGAAGATGATGAATTAATCTGGGCTTACTTGTCTGAACCGGAGATCTTAATGG 1622
DB 481 LeuSerAArgLybMetAArgTyrrTrpAlaAAspAlaAArgThnGlyAAspProAAsnGly 500
QY 1623 AACGACCTGTCTGTGTGGCCAGCTTAAATCTGACAGAGAGTCAAGCTGAGCTTGG 1682
DB 501 GlnGlyAlaProLeuTrpProAlaTyrrThnGlnSerGlnGlnTyrrLeuLybLeu 520
QY 1683 AACATGAGCTCGGACAGACACTCAAAAGAACCGCGGAGTATTTTGGACCGACCATC 1742
DB 521 SerValSerValGlyGlnLybLeuLybGlnGlnValGlnPheTrpMetAAsnHle 540
RESULT 13
US-10-451-168-93
Sequence 93, Application US/10451168
Publication No. US20040091969A1
GENERAL INFORMATION:
APPLICANT: SMITHKLINE BEECHAM CORPORATION
APPLICANT: SMITHKLINE BEECHAM P.L.C.
APPLICANT: GLAXO GROUP LIMITED
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP50039
CURRENT APPLICATION NUMBER: US/10/451,168
PRIOR FILING DATE: 2003-11-12
PRIOR APPLICATION NUMBER: PCT/US01/49232
PRIOR FILING DATE: 2000-12-17
PRIOR APPLICATION NUMBER: 60/256,710
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 60/257,048
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 60/260,482
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: 60/264,922
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/266,797
PRIOR FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 60/276,988
PRIOR FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 60/281,535
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/289,622
PRIOR FILING DATE: 2002-06-28

```

; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-451-168-93

Alignment Scores:
Pred. No.:      1,086-151      Length:      356
Score:          1,780.00      Matches:      340
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     45.97%       Indels:      0
DB:              15           Gaps:        0

US-10-023-515-1 (1-2158) x US-10-451-168-93 (1-356)

QY      819  ATACTGCTCCCATGCGCCAAAGCTTATTCACAAAGCATGAGAGTGGGGTGGCC 878
DB      17  ILeuSerProMetAlaIysGlyLeuPheHisIlybAlaIleMetGluSerGlyValAla 36
QY      879  ATCATCCCTTACCTGAGGCCCATGATTATGAGAAAGTGAAGACCTGCAAGGTGGTGA 938
DB      37  IleIleProTyrLeuGluAlaHisAspTyrGlyIlySerGluAspLeuGlnValValAla 56
QY      939  CATTTCGTGGTAAACATGCGCTCAGACTGTGAGGCCCTGCTGAGGTGCTTGAGACAAA 998
DB      57  HisPheCysGlyAsnAsnAlaSerAspSerGluAlaLeuIleArgCysLeuArgThrIlys 76
QY      999  CCTCCAGAGAGCTGCTGACCCCTGACCCAGAAACAAAGCTTTCACCTGAGTGGTTGAT 1058
DB      77  ProSerIlyGlnLeuLeuThrIleuSerGlnIlySerPheThrArgValValAla 96
QY      1059  GGTGCTTCTTCTTCTATGAGCCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAAT 1118
DB      97  GlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnIlyAlaPheIlybAlaIle 116
QY      1119  CCTTCATCATCTGGAGTCAATACACAGAGTGGCTTCTGCTGCTCTATGAAGAGGT 1178
DB      117  ProSerIleIleGlyValAsnAsnHisGlyCysGlyPheLeuLeuProMetIysGlnAla 136
QY      1179  CCTGAGATCTCAGTGGGCTCCAAACAAGTCCCTTGCCCTCATCTGATACAAACATCCGT 1238
DB      137  ProGluIleLeuSerGlySerAsnIlybSerLeuAlaLeuHisIleuIleGlnMetIleu 156
QY      1239  CACATCCCGCCTCAGTATTGTGACCTTGTGCTAAATGATATCTTCATGACAGCACTCC 1298
DB      157  HisIleProProGlnTyrIleuHisIleuValAlaAsnGlnTyrPheHisAspIlybHisSer 176
QY      1299  CTGACTGAATCCGAGACAGTCTTTCGACTTGTGAGATGTGTTCTTTGTGCTCCT 1358
DB      177  LeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValValPro 196
QY      1359  GCACGTGATCAGACGCTCGATATCAGAGAGTGTGAGACCTGCTACTTCATGAGTTT 1418
DB      197  AlaLeuIleHisAlaArgTyrHisAspAspAlaGlyAlaProValTyrPheTyrGlnPhe 216
QY      1419  CGGACCGCGCCTCAGTGTCTTTGAGACAGCAAGCCGGCTTTTGTCAAGCCGACCACTG 1478
DB      217  ArgHisAspArgProGlnCysPheGlnAspThrIlybProAlaPheValIlybAlaAspHisAla 236
QY      1479  GATGAAGTCCGCTTGTGTGTGGTGTGCTTCTCGAAGGGGACATTTTATGTTCCAA 1538
DB      237  AspGlnValArgPheValPheGlyGlyAlaPheLeuIlybGlyAspIleValMetPheGln 256
QY      1539  GGAGCGACGAGAGAGAGAAAGTACTGAGCGGAGAGATGATAAATPCTGGGCTACCTT 1598
DB      257  GlyAlaThrGlnGlnGlnIlybLeuLeuSerArgIlybMetMetIlybTyrTrpAlaThrPhe 276
QY      1599  GCTCGAACCGGAGATCTTATGAGAAAGCACTGTCTCTGTGCGCAGCTTATATCTGACT 1658
DB      277  AlaArgThrGlnGlyAsnProAsnGlyAsnAspLeuSerLeuThrProAlaTyrAsnLeuThr 296
```

```

QY      1659  GAGCAGTACCTCCAGCTGAGATTGAACATGAGCCTCGGACAGAGACTCAAGAACCGCG 1718
DB      297  GlnGlnTyrLeuGlnMetAspLeuAsnMetSerLeuGlnArgIlybLeuIlybGlnProArg 316
QY      1719  GTGATTTTGGACAGCAACATCCCTGATCCTGTGCTCGGACATGCTCCACACT 1778
DB      317  ValAspPheThrIleThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHisSer 336
QY      1779  CCTCTTCTCTTAACTTTCTCTCTCTCCCGACGCTTTCTTTCTTTGTGCTCCT 1838
DB      337  ProLeuSerSerLeuThrPheLeuSerLeuGlnProPhePhePheCysAlaPro 356

RESULT 14
US-09-925-298-689
; Sequence 689, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P0103
; CURRENT APPLICATION NUMBER: US/09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 689
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-689

Alignment Scores:
Pred. No.:      8,186-102      Length:      549
Score:          1,230.00      Matches:      248
Percent Similarity: 63.84%    Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match:     31.77%       Indels:      24
DB:              9           Gaps:        8

US-10-023-515-1 (1-2158) x US-09-925-298-689 (1-549)

QY      186  GGGCCTTGTGCTGAAGGCGCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCA 245
DB      16  GlyGlnAspSerAlaSerProIleArgThrHisIleGlnGlnValLeuGlySerLeu 35
QY      246  GTCACTGTGCTGGGAAGCCCTGCTGTAACGTTCTCTGGAGTCCCTTTGCTGCT 305
DB      36  ValHisValIlybGlyAlaAsnAlaGlyAlaGlnThrPheLeuGlyIleProPheAlaIlyb 55
QY      306  CCCCCTGGGATCCCTGATTTTACGAACCCGACGCTGATCCGCCCTGGGATTAATCTG 365
DB      56  ProProLeuGlyIlyProLeuArgPheAlaProGlnIlybProGlnIlybSerIlybVal 75
QY      366  CGAGAAGCACTCTTACCTTAATTGTGCTCCAGAAC-----TCGAG 410
DB      76  ArgAspGlyThrThrHisProAlaMetCysLeuGlnAspLeuThrAlaValGlnSerGln 95
QY      411  TGGCTGCTCTTAATGATCAACATGCTCAAGGTCATTACCGGAATTCGGAGTGCAGA 470
DB      96  PheLeu-----SerGlnPheAsnMetThrPheProSerAspSerMetSerGln 111
QY      471  GACTGCTCTACCTGACATGATGATGCGCTGCGCCAGCGGATGACAGGCTCCAACTCCC 530
```

```

Db 112 AapCyLeuTyrLeuSerIleTyrThrProAlaHisSerHisGlySerAsnLeuPro 131
QY 531 GTCTTGATGTTGTTCCAGAGAGTGGCTTCAAGAGTGGTCAAGCTTCCATCTTTGATGGG 590
Db 132 ValMetValTrrP11HisGlyGlyAlaLeuValPheGlyMetAlaSerLeuTyrAspGly 151
QY 591 TCCGCGCTGGTGGCTGATGAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 650
Db 152 SerMetLeuAlaAlaLeuGlnHisValValAlaIleIleGlnTyrArgLeuGlyVal 171
QY 651 TTTGGTTTCTTACACATGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 710
Db 172 LeuGlyPhePheSerThrGlyAspArgHisAlaThrGlyAsnProGlyTyrLeuAspGln 191
QY 711 GTGGTGGTCTGTCTGGTGGTCCAGAAAGACATCGAGTTCTTGGTGGGAGACCCAGCTCT 770
Db 192 ValAlaAlaLeuAlaGlyTrpValGlnGlnHisIleAlaHisPheGlyGlyAsnProAspArg 211
QY 771 GTGACCATCTTTGGCGAGTCCGCGGAGCCATAGTGTCTTACTTATATACGTCTTCCC 830
Db 212 ValThrIlePheGlyGlnUserAlaGlyGlyThrSerValSerSerLeuValValSerPro 231
QY 831 ATGGCCAAAGGCTTATTCACAAAGCCATATGAGAGTGGGAGTGGGAGTGGGAGTGGGAGT 890
Db 232 IleSerGlnGlyLeuPheHisGlyAlaIleMetGlySerGlyValAlaLeuLeuProGly 251
QY 891 CTGAGAGCCCATGATATGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 944
Db 252 LeuIleAla-----SerSerAlaAspValIleSerThrValAlaAlaAsnLeu 267
QY 945 TGTGTGACAAATGCGTCAAGCTTGAAGCCCTGAGTGGTGGTGAAGCAAAACCTTCC 1004
Db 268 SerAlaCysAspGlnValAspSerGlnAlaLeuValGlyCysLeuArgGlyLeuSerLys 287
QY 1005 AAGAGAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGA 1064
Db 288 GlnGlnIleLeuAlaIleAsnLysProPheLysMetIleProGlyValValAspGlyVal 307
QY 1065 TTTCTTCTTAATGAGCTCTAGATTAATGTTGTTGTCAGAAAGCAATTAAGCAATTCCTCC 1124
Db 308 PheLeuProArgHisProGlnGlnLeuAlaSerAlaAspPheGlnProValProSer 327
QY 1125 ATCATGAGAGTCAATAACAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1172
Db 328 IleValGlyValAsnHisAsnGlnLysPheGlyTrrPleIleProLysValMetArgIleTyr 347
QY 1173 GAGGCTCTGAGATCTCTAGTGGTCCAAAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1232
Db 348 AspThrGlnLysGlnMetAspArgGlnLysSerGlnAlaAlaLeuGlnLysMetLeuThr 367
QY 1233 ATCTGACATATCCCGCTCAAGATTAATGACCTTGGTGGTGAATGAATCTTCATGACAAAG 1292
Db 368 LeuLeuMetLeuProThrPheGlyAspLeuLeuAspGlnGlyTrrPleIleGlyAspArgHis 387
QY 1293 CACTCCCTGATGAATCCAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1352
Db 388 GlyAspProGlnThrLeuGlnAlaGlnPheGlnGlnMetMetAlaAspSerMetPheAla 407
QY 1353 GTCCCTGACATGATACAGCTCGATATACAGAGATGCTGGTGAAGTGGTGGTGGTGGTGGTGG 1412
Db 408 IleProAlaLeuGlnValAlaHisPhe---GlnCysSerArgAlaProValTyrPheTyr 426
QY 1413 GAGTTTCCGACACCGGCTCAGTGGTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
Db 427 GlnPheGlnHisGlnProSerTrrPleLysAsnIleArgProProHisMetLysValAsp 446
QY 1473 CAGCTGATGAAGTCCGCTTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1532
Db 447 HisGlyAspGlnLeuProPheValPhe---ArgSerPhePheGlyGlyAsnTrrIleLys 465
QY 1533 TTGCAAGAGAGCAGAGAGAGAGAGATTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1592

```

```

Db 466 Phe-----ThrGlnGlnGlnGlnLeuSerArgLysMetMetLysTrrTrpAla 482
QY 1593 ACCTTGCTTCGAACCGGAGATCTTAATGGGAACAGACTGTCTGTGGCCAGCTTAATAT 1652
Db 483 AsnPheAlaArgAsnGlyAsnProAsnGlyGlnGlyLeuProHisTrrPleLeuPheAsp 502
QY 1653 CTGACTGACAGTACTCAGCTCAGTGGTGGTGAACATGACATGAGCTGGTGGTGGTGGTGGT 1712
Db 503 GlnGlnGlnGlnThrLeuGlnLeuHisAsnGlnProAlaValAlaGlyAlaLeuLysVal 522
QY 1713 CCGCGGTGATTTTGGACAGCAGCAGCATCCCC 1745
Db 523 HisArgLeuGlnPheTrrPleLysValAlaLeuPro 533

RESULT 15
US-10-102-806-689
/ Sequence 689, Application US/10102806
/ Publication No. US20030054421A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: P4103P1C1
/ CURRENT APPLICATION NUMBER: US/10/102,806
/ PRIOR APPLICATION NUMBER: 09/925,298
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05881
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 846
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 689
/ LENGTH: 549
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURES:
/ NAME/KEY: SITE
/ LOCATION: (1)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (7)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-689

Alignment Scores:
Pred. No.: 8,18e-102 Length: 549
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 31.77% Indels: 24
DB: 14 Gaps: 8

US-10-023-515-1 (1-2158) x US-10-102-806-689 (1-549)
QY 186 GGGGCTTCTGCTGAAGGAGCAGAGAAACACAGCTGGAGTGAATTCAGGAGCAAGCA 245
Db 16 GlyGlnAspSerAlaSerProIleArgThrHisIsthrrGlnValLeuGlySerLeu 35
QY 246 GTCACTGTGCTGGGAAAGCCCTGTGCTGTGAACGTGTTCTCGAAGTCCCTTGTGCT 305
Db 36 ValHisValLysGlyAlaAsnAlaGlyAlaGlnThrPheLeuGlyIleProPheAlaLys 55
QY 306 CCGCGCTGGGATCCCTGCAATTAACGAACCGGAGCTGGATGCGCTGGGAGTAATCTTG 365
Db 56 ProPheLeuGlyProLeuArgPheAlaProGlnProProGlnUserTrrPserGlyVal 75
QY 366 CGAAGAGCAGCTCTTACCTTAATTTGTGCTCGAAGAC-----TCAGAG 410
Db 76 ArgAspGlyTrrHisPheAlaMetCysLeuGlnAspLeuThrAlaValGlnSerGln 95
QY 411 TGGTGTCTTAAGTCAACACATGCTCAAGTGAATTCAGGAGTGGTGGTGGTGGTGGTGGTGGTGG 470

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 22:52:56 ; Search time 1123.45 Seconds
(without alignments)
9634.209 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 1746

Sequence: 1 atgcacagcgactactctc.....ttctcttctgtgctctctga 1746

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 309530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
2	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
3	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
4	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
5	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appli
6	1652.8	94.7	1728	US-10-451-168-40	Sequence 45, Appli
7	1651.6	94.6	2232	US-10-433-256-23	Sequence 23, Appli

8	1631.2	93.4	2092	US-10-104-047-249	Sequence 249, App
9	1435.6	82.0	1746	US-10-114-270-195	Sequence 195, App
10	1432.4	82.0	1746	US-10-451-168-41	Sequence 41, Appli
11	1386.2	75.4	1857	US-10-381-898-20	Sequence 20, Appli
12	1344.4	77.0	1962	US-10-094-749-736	Sequence 736, App
13	1099	62.9	2145	US-10-233-933A-3	Sequence 3, Appli
14	1063.2	60.9	1629	US-10-233-933A-1	Sequence 1, Appli
15	1021.4	58.5	1071	US-10-451-168-42	Sequence 42, App
16	508.6	29.1	965	US-10-302-172-155	Sequence 155, App
17	406.6	22.3	2191	US-09-880-107-3854	Sequence 3854, Ap
18	406.6	22.3	2191	US-09-954-531-1038	Sequence 1038, Ap
19	406.6	22.3	2191	US-09-873-367C-828	Sequence 828, App
20	406.6	22.3	2191	US-09-873-367C-829	Sequence 829, App
21	406.6	22.3	2191	US-10-858-271-27	Sequence 27, Appli
22	406.6	22.3	2191	US-10-843-641A-828	Sequence 828, App
23	406.6	22.3	2191	US-10-843-641A-829	Sequence 829, App
24	406.6	22.3	2191	US-10-843-641A-2105	Sequence 2105, Ap
25	406.6	22.3	2484	US-09-925-298-271	Sequence 271, App
26	406.6	22.3	2484	US-10-102-806-471	Sequence 471, App
27	400.2	22.9	2092	US-10-956-157-4153	Sequence 4153, Ap
28	400.2	22.9	3824	US-09-931-836-322	Sequence 322, Appli
29	400.2	22.9	3824	US-10-036-342-22	Sequence 22, Appli
30	400.2	22.9	3824	US-10-036-041-22	Sequence 22, Appli
31	400.2	22.9	3824	US-10-028-072-541	Sequence 541, App
32	400.2	22.9	3824	US-10-035-855-22	Sequence 22, Appli
33	400.2	22.9	3824	US-10-140-808-541	Sequence 541, App
34	400.2	22.9	3824	US-10-121-049-541	Sequence 541, App
35	400.2	22.9	3824	US-10-123-904-541	Sequence 541, App
36	400.2	22.9	3824	US-10-140-470-541	Sequence 541, App
37	400.2	22.9	3824	US-10-175-746-541	Sequence 541, App
38	400.2	22.9	3824	US-10-176-918-541	Sequence 541, App
39	400.2	22.9	3824	US-10-176-918-541	Sequence 541, App
40	400.2	22.9	3824	US-10-227-864-209	Sequence 209, App
41	400.2	22.9	3824	US-10-036-214-22	Sequence 22, Appli
42	400.2	22.9	3824	US-10-137-865-541	Sequence 541, App
43	400.2	22.9	3824	US-10-140-474-541	Sequence 541, App
44	400.2	22.9	3824	US-10-035-719-22	Sequence 22, Appli
45	400.2	22.9	3824	US-10-142-431-541	Sequence 541, App

ALIGNMENTS

RESULT 1
US-10-023-515-3
; Sequence 3, Application US/10023515
; Publication No. US20020182636A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023, 515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279, 508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-515-3

Query Match 100.0%; Score 1746; DB 13; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGCACAGGAGCTACTCTCATCTCTTCACATGATGCTTTTCTCTGATTCACAGCCC 60
|||||

Db 1 ATGCCACAGGGAGCTTACTTCACTGCTTCACAATGAGTGTCTTTCTGATCTTCAGAGCC 60
Qy 61 CTGTTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGGCCCAAGAGAAC 120
Db 61 CTGTTGGGACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGGCCCAAGAGAAC 120
Qy 121 ACCAGGCTGGATGAGATTCAAGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTG 180
Db 121 ACCAGGCTGGATGAGATTCAAGGCAAGCAAGTCACTGTGTGGGAAGCCCTGTGCTGTG 180
Qy 181 AACGGTCTCTGGAGTCCCTTTGCTGCTCTCCCGCTGGAGTCCCTGCAATTTACGAAC 240
Db 181 AACGGTCTCTGGAGTCCCTTTGCTGCTCTCCCGCTGGAGTCCCTGCAATTTACGAAC 240
Qy 241 CCGAGCCTGATGCTCCCTGGGATTAATTGGAGAGGCACTCTCTACCTTAATTTGTG 300
Db 241 CCGAGCCTGATGCTCCCTGGGATTAATTGGAGAGGCACTCTCTACCTTAATTTGTG 300
Qy 301 CTCGAGAACTCAGAGTGTGCTCTTATGATCAACATGCTCAAGGTGCAATTAACCGAAA 360
Db 301 CTCGAGAACTCAGAGTGTGCTCTTATGATCAACATGCTCAAGGTGCAATTAACCGAAA 360
Qy 361 TTGGAGATGTCAGAGACTGCTCTACCTGAACAATCTATGCGCTGCGCCACGCCGATACA 420
Db 361 TTGGAGATGTCAGAGACTGCTCTACCTGAACAATCTATGCGCTGCGCCACGCCGATACA 420
Qy 421 GGTCTCCAGCTCCCGCTTGTGTGTGTGCTCCAGAGAGTGCCTTCAAGACTGTGCTAGCC 480
Db 421 GGTCTCCAGCTCCCGCTTGTGTGTGTGCTCCAGAGAGTGCCTTCAAGACTGTGCTAGCC 480
Qy 481 TCCATCTTTGATGGATCCGCTGCTGCTGCTATGAGAGAGCTGCTGTTGTGTGCTCAG 540
Db 481 TCCATCTTTGATGGATCCGCTGCTGCTGCTATGAGAGAGCTGCTGTTGTGTGCTCAG 540
Qy 541 TACCGGCTAGGAATATTGTTTCTTCAACAATGAGGATCAAGATGCTCCGGGAACTGG 600
Db 541 TACCGGCTAGGAATATTGTTTCTTCAACAATGAGGATCAAGATGCTCCGGGAACTGG 600
Qy 601 GCGTTAAAGAACAGGTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 GCGTTAAAGAACAGGTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 GGGAGACCCAGCTGTGTGACATCTTTGAGGAGTCCGCGGGAGCCATAGTGTCTTCTAGT 720
Db 661 GGGAGACCCAGCTGTGTGACATCTTTGAGGAGTCCGCGGGAGCCATAGTGTCTTCTAGT 720
Qy 721 CTTATTACTGTCTCCATGAGCCAAAGCTTATTCACAAAGCCATGAGAGTGGGGTG 780
Db 721 CTTATTACTGTCTCCATGAGCCAAAGCTTATTCACAAAGCCATGAGAGTGGGGTG 780
Qy 781 GGCATCATCTCCCTTACCTGAGGCGCATGATTAAGAAGAGTGGAGACTTCAAGTGTG 840
Db 781 GGCATCATCTCCCTTACCTGAGGCGCATGATTAAGAAGAGTGGAGACTTCAAGTGTG 840
Qy 841 GCACATTTCTGTGTAAACAATGCGTCAAGCTCTGAGGCGCTGTGAGGTGCTTGAAGCA 900
Db 841 GCACATTTCTGTGTAAACAATGCGTCAAGCTCTGAGGCGCTGTGAGGTGCTTGAAGCA 900
Qy 901 AAACCTTCAAGAGCTGTGCTGAGCCCTCAAGCCAGCAAAACAAAGTCTTCACTCGAGTGT 960
Db 901 AAACCTTCAAGAGCTGTGCTGAGCCCTCAAGCCAGCAAAACAAAGTCTTCACTCGAGTGT 960
Qy 961 GATGATGCTTCTTCTCTAATGAGCTCTTATGATCTTATGCTCAAGAAAGCAATTTAAAGCA 1020
Db 961 GATGATGCTTCTTCTCTAATGAGCTCTTATGATCTTATGCTCAAGAAAGCAATTTAAAGCA 1020
Qy 1021 ATTCTCTTCATCATGAGATCAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAG 1080
Db 1021 ATTCTCTTCATCATGAGATCAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAG 1080
Qy 1081 GCTCTGAGATCTCAGTGTGCTCCAAAGTCCCTTCCATCTGATCAAAACATC 1140
Db 1081 GCTCTGAGATCTCAGTGTGCTCCAAAGTCCCTTCCATCTGATCAAAACATC 1140

Qy 1141 CTCACATCCCGCTGAGTATTGACCTTGTGGCTTAATGAATATCTTCATGACAGAC 1200
Db 1141 CTCACATCCCGCTGAGTATTGACCTTGTGGCTTAATGAATATCTTCATGACAGAC 1200
Qy 1201 TCCCTGACTGAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTTCTTTGTGCTC 1260
Db 1201 TCCCTGACTGAATCCGAGACAGTCTTGTGACTTGTGAGATGTGTTCTTTGTGCTC 1260
Qy 1261 CCTGACATGATCAAGCTGATATCAAGAGTGTGTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 CCTGACATGATCAAGCTGATATCAAGAGTGTGTGCTGCTGCTGCTGCTGCTGCTG 1320
Qy 1321 TTTGGACACCGGCTCAGTCTTGTGAAGACAGAGCGGCTTTGTCAAGCCGACAC 1380
Db 1321 TTTGGACACCGGCTCAGTCTTGTGAAGACAGAGCGGCTTTGTCAAGCCGACAC 1380
Qy 1381 GCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Db 1381 GCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
Qy 1441 GAGGAGCCAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAG 1500
Db 1441 GAGGAGCCAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAG 1500
Qy 1501 TTTGCTGCAACCGGGAATCTTAATGGGAACGACTGTCTGTGTGCGCAGCTTATATCTG 1560
Db 1501 TTTGCTGCAACCGGGAATCTTAATGGGAACGACTGTCTGTGTGCGCAGCTTATATCTG 1560
Qy 1561 ACTGAGAGTCTCTCAAGCTGAGTGAACATGAGCCTCGGACAGAGACTCAAGAACCG 1620
Db 1561 ACTGAGAGTCTCTCAAGCTGAGTGAACATGAGCCTCGGACAGAGACTCAAGAACCG 1620
Qy 1621 CGGTGATTTTGTGACCAAGCAACATCCCTGATCTGTCTGCTGCTGCTGCTGCTGCT 1680
Db 1621 CGGTGATTTTGTGACCAAGCAACATCCCTGATCTGTCTGCTGCTGCTGCTGCTGCT 1680
Qy 1681 AGTCT 1740
Db 1681 AGTCT 1740
Qy 1741 CTTTGA 1746
Db 1741 CTTTGA 1746

RESULT 2
US-10-674-636-3
Sequence 3, Application US/10674636
Publication No. US20040086922A1
GENERAL INFORMATION:
APPLICANT: Curles, Roy A. J.
TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
FILE REFERENCE: 10448-122001
CURRENT APPLICATION NUMBER: US/10/674, 636
PRIOR FILING DATE: 2003-09-29
PRIOR APPLICATION NUMBER: US/10/023, 515
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: 60/256, 369
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/279, 508
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1746
TYPE: DNA
ORGANISM: Homo sapiens
US-10-674-636-3
Query Match 100.0%; Score 1746; DB 18; Length 1746;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Indels 0; Gaps 0;

Oy	1	ATGCAAGAGGAACTTACTTCAATCTGCTTCAAAATGAGCTTTTCTGATTTCTCAGAGCC	60
Db	1	ATGCCAAGAGGAACTTACTTCAATCTGCTTCAAAATGAGCTTTTCTGATTTCTCAGAGCC	60
Oy	61	CTGTGTGGACA CAGACAGTGGGGAAAAA CTGGGCTTGTGCTGAAGGGCCA CAGAGAAC	120
Db	61	CTGTGTGGGACA CAGACAGTGGGGAAAAA CTGGGCTTGTGCTGAAGGGCCA CAGAGAAC	120
Oy	121	ACCAAGGCTGGGATGGATT CAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGTGCTGTG	180
Db	121	ACCAAGGCTGGGATGGATT CAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGTGCTGTG	180
Oy	181	AACGTTCTCTCGGAGTCCCCTTTGCTGCTCCCGGCTGGGATCCCTGCAATTTACGAAC	240
Db	181	AACGTTCTCTCGGAGTCCCCTTTGCTGCTCCCGGCTGGGATCCCTGCAATTTACGAAC	240
Oy	241	CCGAGCCCTGGATCGCCCTGGGATTAATCTTGGAGAAAGCCACCTCTCAACTTAATTGTGC	300
Db	241	CCGAGCCCTGGATCGCCCTGGGATTAATCTTGGAGAAAGCCACCTCTCAACTTAATTGTGC	300
Oy	301	CTCCAGAACTCAAGATGAGCTGTCTTAATATCAACAATGCTCAAGGTGCATTAACCCGAAA	360
Db	301	CTCCAGAACTCAAGATGAGCTGTCTTAATATCAACAATGCTCAAGGTGCATTAACCCGAAA	360
Oy	361	TTTGGAGGTGAGAAAGATGACTGACTACCTGAACATCTAATGGCGCTGCCACAGCCGATACA	420
Db	361	TTTGGAGGTGAGAAAGATGACTGACTACCTGAACATCTAATGGCGCTGCCACAGCCGATACA	420
Oy	421	GAGCTCCAAGCTCCCCTGTCTTGATGTGTCTCCAGAGAGGTGCCTTCAAGACTGGCTCAGCC	480
Db	421	GAGCTCCAAGCTCCCCTGTCTTGATGTGTCTCCAGAGAGGTGCCTTCAAGACTGGCTCAGCC	480
Oy	481	TTCATCTTTTAATGGGTCCGCCCTGGCTGCTGCTATAGAGACGTGTGATGTGTGCTCCAG	540
Db	481	TTCATCTTTTAATGGGTCCGCCCTGGCTGCTGCTATAGAGACGTGTGATGTGTGCTCCAG	540
Oy	541	TACCGGGCTAGGAATATTGTTGTTTCTTCAACAATAGGAATCAGCAATGCTCCGGGGAAACTGG	600
Db	541	TACCGGGCTAGGAATATTGTTGTTTCTTCAACAATAGGAATCAGCAATGCTCCGGGGAAACTGG	600
Oy	601	GCTTTCAAAGACCAAGTGGCTGTCTGTCTTGAGTCCAGAAACATCGAATTTCTTCGGT	660
Db	601	GCTTTCAAAGACCAAGTGGCTGTCTGTCTTGAGTCCAGAAACATCGAATTTCTTCGGT	660
Oy	661	GGGGACCCCAAGCTCTGTGACATCTTTGGCAGGTCCGGGAGGCATTAAGTGTCTTAGT	720
Db	661	GGGGACCCCAAGCTCTGTGACATCTTTGGCAGGTCCGGGAGGCATTAAGTGTCTTAGT	720
Oy	721	CTTAATACGTGTCCCAATGGCCAAAGGCTTAATTCACAAGAGCAATCATAGGAGTGGGGTGG	780
Db	721	CTTAATACGTGTCCCAATGGCCAAAGGCTTAATTCACAAGAGCAATCATAGGAGTGGGGTGG	780
Oy	781	GCCATCATCTCCCTTAATCTGAGAGGCCCATGATTAATGAGAAGATGAGACCTTGCAGGTGTT	840
Db	781	GCCATCATCTCCCTTAATCTGAGAGGCCCATGATTAATGAGAAGATGAGACCTTGCAGGTGTT	840
Oy	841	GCAATTTCTGTGTAAACAATGCGTCAGACTCTGAGGCCCTGTGAGGTGCTTGAGACA	900
Db	841	GCAATTTCTGTGTAAACAATGCGTCAGACTCTGAGGCCCTGTGAGGTGCTTGAGACA	900
Oy	901	AAACCCCTCAAGAGCTGTGACCTGACCTGACGCAAAAA CAAAGTCTTTTCACTCGAGTGGTT	960
Db	901	AAACCCCTCAAGAGCTGTGACCTGACCTGACGCAAAAA CAAAGTCTTTTCACTCGAGTGGTT	960
Oy	961	GATGGTGTCTTTCTTCTCTAATGAGCTCTAGATCTAATGTTCTCAAGAAAGCAATTTAAACA	1020
Db	961	GATGGTGTCTTTCTTCTCTAATGAGCTCTAGATCTAATGTTCTCAAGAAAGCAATTTAAACA	1020
Oy	1021	ATTGCTTTCATCATGGAAGTCAATTAACAAGATGTGGCTTCTGTGCTGCTATGAGAG	1080
Db	1021	ATTGCTTTCATCATGGAAGTCAATTAACAAGATGTGGCTTCTGTGCTGCTATGAGAG	1080

Dp	102.	ATTCTTCACATCANTCGGAGTCAATPACACGAGTGTGGTTCCTGCTGCTCATGAAAGG	1080
Oy	1081	GCTCTCTGAGATCCTCAATGGGCTCCACAAAGTCCCTTGCCCTCCATCTGATATPCAAACATC	1140
Dp	1081	GCTCTCTGAGATCCTCAATGGGCTCCACAAAGTCCCTTGCCCTCCATCTGATATPCAAACATC	1140
Oy	1141	CTGCACATCCCGGCTCAGTATTTTGGCACCTTGTTGGCTAATGATATPACTTCCATGACAAACAC	1200
Dp	1141	CTGCACATCCCGGCTCAGTATTTTGGCACCTTGTTGGCTAATGATATPACTTCCATGACAAACAC	1200
Oy	1201	TCCCTGACGTGAATCCGAGACAGTCTTCTGGACCTTGTGGAGATGTGTTCTTTGTGATC	1260
Dp	1201	TCCCTGACGTGAATCCGAGACAGTCTTCTGGACCTTGTGGAGATGTGTTCTTTGTGATC	1260
Oy	1261	CTTGCACTGATCACAAGCTCGATATCACAAGATGCTGTGACACTGTCTACTTCTTATGAG	1320
Dp	1261	CTTGCACTGATCACAAGCTCGATATCACAAGATGCTGTGACACTGTCTACTTCTTATGAG	1320
Oy	1321	TTTTCGGGACCGGGCTCAGTGTCTTGAAGACACGAAAGCGGCTTTGTCAAGCCGACAC	1380
Dp	1321	TTTTCGGGACCGGGCTCAGTGTCTTGAAGACACGAAAGCGGCTTTGTCAAGCCGACAC	1380
Oy	1381	GCTGATGAAGTCCGCTTTGTGTTGGTGTGCTTCCCTGAAGGGGGACAATGTTATGTTTC	1440
Dp	1381	GCTGATGAAGTCCGCTTTGTGTTGGTGTGCTTCCCTGAAGGGGGACAATGTTATGTTTC	1440
Oy	1441	GAAGGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTAAC	1500
Dp	1441	GAAGGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTAAC	1500
Oy	1501	TTTGTCTGGAACCGGGAATCCTTAATGGGAACGACTGTCTCTGTGGCCAGCTTATTAATCTG	1560
Dp	1501	TTTGTCTGGAACCGGGAATCCTTAATGGGAACGACTGTCTCTGTGGCCAGCTTATTAATCTG	1560
Oy	1561	ACTGAGCAGTACCTCCAGCTGACCTTGAACATGAGCTCTGGGACAGAGACTCAAAAGAACCG	1620
Dp	1561	ACTGAGCAGTACCTCCAGCTGACCTTGAACATGAGCTCTGGGACAGAGACTCAAAAGAACCG	1620
Oy	1621	CGGGTGAGATTTTGGACACAGACATCCCTCTGATCTGTCTGTCCGACATGCTCCAC	1680
Dp	1621	CGGGTGAGATTTTGGACACAGACATCCCTCTGATCTGTCTGTCCGACATGCTCCAC	1680
Oy	1681	AGTCTCTTTTCTCTTAACTTTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCT	1740
Dp	1681	AGTCTCTTTTCTCTTAACTTTCTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGCT	1740
Oy	1741	CCCTTGA 1746	
Dp	1741	CCCTTGA 1746	

```

1 RESULT 3
2 US-10-023-515-1
3 Sequence 1, Application US/10023515
4 Publication No. US20020182636A1
5 GENERAL INFORMATION:
6 APPLICANT: Curtis, Rory A. J.
7 APPLICANT: Sllas-Santiago, Immaculada
8 TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL-ESTERASE
9 TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
10 FILE REFERENCE: 10046-122001
11 CURRENT APPLICATION NUMBER: US/10/023,515
12 CURRENT FILING DATE: 2001-12-18
13 PRIOR APPLICATION NUMBER: 60/256,369
14 PRIOR FILING DATE: 2000-12-18
15 PRIOR APPLICATION NUMBER: 60/279,508
16 PRIOR FILING DATE: 2001-03-28
17 NUMBER OF SEQ ID NOS: 6
18 SOFTWARE: FastSeq for Windows Version 4.0
19 SEQ ID NO 1
20 LENGTH: 2158
21 TYPE: DNA
22 ORGANISM: Homo sapiens

```

FEATURE:
; NAME/KEY: CDS
; LOCATION: (96) ... (1838)
US-10-023-515-1

Query Match 100.0%; Score 1746; DB 13; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCCACAGGGAATTACTTCTGCTTCAACAGTGTGCTTTTCTGATTTCTCCAGCC 60
Db 96 ATGCCACAGGGAATTACTTCTGCTTCAACAGTGTGCTTTTCTGATTTCTCCAGCC 60
QY 61 CTGTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAAGGACCAAGAGAAC 120
Db 156 CTGTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAAGGACCAAGAGAAC 120
QY 121 ACCAGGCTGGAGTGGATTTCAAGGCAAGCAAGTCACTGTGCTGGAGAACCTGTGCTGTG 180
Db 216 ACCAGGCTGGAGTGGATTTCAAGGCAAGCAAGTCACTGTGCTGGAGAACCTGTGCTGTG 275
QY 181 AACGTGTCTCGGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGGGATTTAGAAC 240
Db 276 AACGTGTCTCGGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGGGATTTAGAAC 335
QY 241 CCGGAGCTTGATCGCCCTGGGATTAATTGGAGAACCACTCTTACCTTAATTTGTGC 300
Db 336 CCGGAGCTTGATCGCCCTGGGATTAATTGGAGAACCACTCTTACCTTAATTTGTGC 395
QY 301 CTCAGAACTCAGAGTGGCTGCTCTTATGATCAACATGCTGAAGTGAATTAACCGAAA 360
Db 396 CTCAGAACTCAGAGTGGCTGCTCTTATGATCAACATGCTGAAGTGAATTAACCGAAA 455
QY 361 TTGGAGTGTCAAGAACAGTGCCTTAACCTGAATGAGCGCTGCCACCGCATACA 420
Db 456 TTGGAGTGTCAAGAACAGTGCCTTAACCTGAATGAGCGCTGCCACCGCATACA 515
QY 421 GGCTCAACCTCCCGCTTTGGTGTGCTCCAGAGGTGCTTTCAAGCTGAGCTCAGCC 480
Db 516 GGCTCAACCTCCCGCTTTGGTGTGCTCCAGAGGTGCTTTCAAGCTGAGCTCAGCC 575
QY 481 TCCATCTTTGATGGTCCGCTGGCTGCTGCTATGAGGAGCTGCTGCTGCTGCTGCTG 540
Db 576 TCCATCTTTGATGGTCCGCTGGCTGCTGCTATGAGGAGCTGCTGCTGCTGCTGCTG 635
QY 541 TACCGGCTAGAGAAATTTGGTCTTCAACACATGAGTCAAGTCAAGTCAAGTCAAG 600
Db 636 TACCGGCTAGAGAAATTTGGTCTTCAACACATGAGTCAAGTCAAGTCAAGTCAAG 695
QY 601 GCCTTCAAGAACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 696 GCCTTCAAGAACAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
QY 661 GGGGACCCCAAGCTGTGACCATCTTTGGGAGTCCGCGGAGGACATTAAGTTCTAGT 720
Db 756 GGGGACCCCAAGCTGTGACCATCTTTGGGAGTCCGCGGAGGACATTAAGTTCTAGT 815
QY 721 CTATATCTGTCTCCATGAGGCAAGGCTTATTCACAAAGCATATGAGAGTGGGCTG 780
Db 816 CTATATCTGTCTCCATGAGGCAAGGCTTATTCACAAAGCATATGAGAGTGGGCTG 875
QY 781 GCATCATCTCTTAAGTGGAGGCCATGATTAAGAGAGTGAAGAGTGAAGTGAAGT 840
Db 876 GCATCATCTCTTAAGTGGAGGCCATGATTAAGAGAGTGAAGAGTGAAGTGAAGT 935
QY 841 GGCATTTTGTGTGAACATGCTGAGACTGTGAGGCTGCTGAGTGGCTGAGAGACA 900
Db 936 GGCATTTTGTGTGAACATGCTGAGACTGTGAGGCTGCTGAGTGGCTGAGAGACA 995
QY 901 AAACCTCTCAAGAGCTGCTGAGCCCTGAGGCAAGAAACAAAGCTTTCACTGAGTGT 960
Db 996 AAACCTCTCAAGAGCTGCTGAGCCCTGAGGCAAGAAACAAAGCTTTCACTGAGTGT 1055
```

```
QY 961 GATGTGCTTTCTTCTTATGAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAGCA 1020
Db 1056 GATGTGCTTTCTTCTTATGAGCTCTAGATCTATTGTCTCAGAAAGCATTTAAGCA 1115
QY 1021 ATTCTTCATCATATGAGGATCAATAACCAAGTGTGCTTCTGCTCTATGAGAG 1080
Db 1116 ATTCTTCATCATATGAGGATCAATAACCAAGTGTGCTTCTGCTCTATGAGAG 1175
QY 1081 GCTCTGAGATCTGAGTGTCTCAACAAAGTCCCTTGGCTCCATCTGATCAAAACATC 1140
Db 1176 GCTCTGAGATCTGAGTGTCTCAACAAAGTCCCTTGGCTCCATCTGATCAAAACATC 1235
QY 1141 CTGCAATCTCCGCTCAAGTATTGCACTTGTGCTAATGAATATTCCATGACAGAC 1200
Db 1236 CTGCAATCTCCGCTCAAGTATTGCACTTGTGCTAATGAATATTCCATGACAGAC 1295
QY 1201 TCCCTGATGAAATCCGAGACAGTCTTCTGAGACTTGTGAGATGTGTTCTTGTGCTC 1260
Db 1296 TCCCTGATGAAATCCGAGACAGTCTTCTGAGACTTGTGAGATGTGTTCTTGTGCTC 1355
QY 1261 CCTGACTGATCAAGCTCGATATCAAGAGATCTGATGACCTGTCTACTTATGAG 1320
Db 1356 CCTGACTGATCAAGCTCGATATCAAGAGATCTGATGACCTGTCTACTTATGAG 1415
QY 1321 TTGGGACACGGGCTCAAGTGTGTTGAAGACAGAAAGCCGCTTTGTCAAGCCGACAC 1380
Db 1416 TTGGGACACGGGCTCAAGTGTGTTGAAGACAGAAAGCCGCTTTGTCAAGCCGACAC 1475
QY 1381 GCTGATGAGTCCGCTTTGTTGTGCTGCTGCTCTGCTGAAAGGGGACATGTTATGTC 1440
Db 1476 GCTGATGAGTCCGCTTTGTTGTGCTGCTGCTCTGCTGAAAGGGGACATGTTATGTC 1535
QY 1441 GAAGGACCAAGAGAGGAGAAAGTTAAGTGAAGCTGAGCGGAAATGATGAATACTGGCTAAC 1500
Db 1536 GAAGGACCAAGAGAGGAGAAAGTTAAGTGAAGCTGAGCGGAAATGATGAATACTGGCTAAC 1595
QY 1501 TTGCTCGAACCAGGAAATCTTAATGGAAGACCTGCTCTGAGGCACTTAATATCTG 1560
Db 1596 TTGCTCGAACCAGGAAATCTTAATGGAAGACCTGCTCTGAGGCACTTAATATCTG 1655
QY 1561 ACTGAGAGTACTCTCAAGCTGACCTTGAACATGAGGCTGAGACAGAGCTCAAGAACCG 1620
Db 1656 ACTGAGAGTACTCTCAAGCTGACCTTGAACATGAGGCTGAGACAGAGCTCAAGAACCG 1715
QY 1621 CGGAGGATTTTGGACAGACCAATCCCTGATCTGTGCTCGACATGCTCCAC 1680
Db 1716 CGGAGGATTTTGGACAGACCAATCCCTGATCTGTGCTCGACATGCTCCAC 1775
QY 1681 AGTCTCTTCTCTTAATCTTCTCTCTCTCTCAAGCTTTCTTTTCTTGTGCT 1740
Db 1776 AGTCTCTTCTCTTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
QY 1741 CCTTGA 1746
Db 1836 CCTTGA 1841
```

RESULT 4
US-10-674-636-1
; Sequence 1, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674, 636
; CURRENT FILING DATE: 2003-09-29
; PRIOR APPLICATION NUMBER: US/10/023, 515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: 60/279,508
PRIOR FILING DATE: 2001-03-28
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2158
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (96)...(1838)
US-10-674-636-1

Query Match 100.0%; Score 1746; DB 18; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCCAGAGGACTTACTTCACTCTTCACTGATGCTTTTCTGATTTCTCAGCCC 60
96 ATGCCAGAGGACTTACTTCACTCTTCACTGATGCTTTTCTGATTTCTCAGCCC 155
61 CTGTTGGGACACAGCAGTGGGAAAACTGGCCCTTCTGCTGAAGGGCCACAGAGAAC 120
156 CTGTTGGGACACAGCAGTGGGAAAACTGGCCCTTCTGCTGAAGGGCCACAGAGAAC 215
121 ACCAGGCTGGATGATTCAGGGGCAAGCAATGCTGCTGGGAAAGCCCTGGCTGTG 180
216 ACCAGGCTGGATGATTCAGGGGCAAGCAATGCTGCTGGGAAAGCCCTGGCTGTG 275
181 AACGTTCTCTGAGATCCCTTTTCTGCTGCTCCCGCTGGATCCCTGGAATTTACGAA 240
276 AACGTTCTCTGAGATCCCTTTTCTGCTGCTCCCGCTGGATCCCTGGAATTTACGAA 335
241 CCGCAGCTGATCCCTCTGGGATTAATTTGGGAAAGCACTTCTCACTTAATTTGTC 300
336 CCGCAGCTGATCCCTCTGGGATTAATTTGGGAAAGCACTTCTCACTTAATTTGTC 395
301 CTCGAGACTGAGATGCTGCTTTAGATCAACATGCTCAAGATGATTAACCGGAA 360
396 CTCGAGACTGAGATGCTGCTTTAGATCAACATGCTCAAGATGATTAACCGGAA 455
361 TTCGAGATGAGAGACTGCTTACCTGAACATTTAGCGCTGCCACCGCGATACA 420
456 TTCGAGATGAGAGACTGCTTACCTGAACATTTAGCGCTGCCACCGCGATACA 515
421 GGCCTCAAGCTCCCGCTTGTGTGTTCCTGAGAGGCTTCAAGACTGGCTCAGCC 480
516 GGCCTCAAGCTCCCGCTTGTGTGTTCCTGAGAGGCTTCAAGACTGGCTCAGCC 575
481 TCCATCTTGAATGGGTCGGCCCTGCTGCTATGAGAGCGTGTGTGTGTGTGTG 540
576 TCCATCTTGAATGGGTCGGCCCTGCTGCTATGAGAGCGTGTGTGTGTGTGTG 635
541 TACCGGCTAGGAATTTTGTCTTTCACCAATGAGGATCAGCATGCTCCGGGAACTGG 600
636 TACCGGCTAGGAATTTTGTCTTTCACCAATGAGGATCAGCATGCTCCGGGAACTGG 695
601 GCTTCAAGAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
696 GCTTCAAGAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
661 GGGGAGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
756 GGGGAGCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
721 CTATATCTGCTCCATGAGGCAAGAGCTTATTCACAAAGCATGATGAGAGTGGGGTG 780
816 CTATATCTGCTCCATGAGGCAAGAGCTTATTCACAAAGCATGATGAGAGTGGGGTG 875
781 GGCATCATCTCTTACCTGAGGCGCAATGATTTGAGAAAGTGAAGACTCTGAGGTGTT 840
876 GGCATCATCTCTTACCTGAGGCGCAATGATTTGAGAAAGTGAAGACTCTGAGGTGTT 935

841 GCACATTTCTGTGTGAACATGCGTCAGACTCTGAGGCCCTGCTGAGGTGCTGAGACA 900
936 GCACATTTCTGTGTGAACATGCGTCAGACTCTGAGGCCCTGCTGAGGTGCTGAGACA 995
901 AATCCCTCAAGAGAGCTGCTGAGCCCTCAGCCAGAAACAAAGTCTTTCACTGAGTGT 960
996 AATCCCTCAAGAGAGCTGCTGAGCCCTCAGCCAGAAACAAAGTCTTTCACTGAGTGT 1055
961 GATGTTGCTTTCTTTCTTAATGAGCCTTAATGAGTCTTAATGAGTCTTAATGAGTCT 1020
1056 GATGTTGCTTTCTTTCTTAATGAGCCTTAATGAGTCTTAATGAGTCTTAATGAGTCT 1115
1021 ATTCTTCATCATGAGAGTCAATTAACAGAGAGTGTGCTTCTGCTGCTATGAGAGAG 1080
1116 ATTCTTCATCATGAGAGTCAATTAACAGAGAGTGTGCTTCTGCTGCTATGAGAGAG 1175
1081 GCTCTGAGATCTCTGAGTGTCTCAACAGATCCCTTCCCTTCATCTGATTAACAAATC 1140
1176 GCTCTGAGATCTCTGAGTGTCTCAACAGATCCCTTCCCTTCATCTGATTAACAAATC 1235
1141 CTGCAATCCCGGCTCAATTTTGAACCTTGTGCTAATGAATCTTCATGAGAGAGAC 1200
1236 CTGCAATCCCGGCTCAATTTTGAACCTTGTGCTAATGAATCTTCATGAGAGAGAC 1295
1201 TCCCTGACTGAATTCGAGAGAGTCTTCTGAGCTTCTGAGATGTCTTTGTGTGTC 1260
1236 TCCCTGACTGAATTCGAGAGAGTCTTCTGAGCTTCTGAGATGTCTTTGTGTGTC 1355
1261 CTTGCACTGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTACTTCTATGAG 1320
1356 CTTGCACTGATCAAGCTGATATCAAGAGATGCTGTGCACTGTCTACTTCTATGAG 1415
1321 TTTGGGACCGGCTCAATGCTTTGAAGAGAGAGCGGCTTTTGTGAAGCGGACAC 1380
1416 TTTGGGACCGGCTCAATGCTTTGAAGAGAGAGCGGCTTTTGTGAAGCGGACAC 1475
1381 GCTGATGAAGTCCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
1476 GCTGATGAAGTCCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1535
1441 GAGGAGCCAGAGAGAGAGAGATTAATGAGCCGGAAGATGATGAATTAATGAGCTAAC 1500
1536 GAGGAGCCAGAGAGAGAGAGATTAATGAGCCGGAAGATGATGAATTAATGAGCTAAC 1595
1501 TTTGCTGAAACCGGGAATCTTAATGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
1596 TTTGCTGAAACCGGGAATCTTAATGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1655
1561 ACTGAGAGTACCTTCAGCTGAGCTTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1620
1656 ACTGAGAGTACCTTCAGCTGAGCTTGAACATGAGCTCGGACAGAGACTCAAGAACCG 1715
1621 CGGGTGAATTTTGGACAGAGACATCCCTGATCTGTCTGCTGCTGCTGCTGCTGCTGCT 1680
1716 CGGGTGAATTTTGGACAGAGACATCCCTGATCTGTCTGCTGCTGCTGCTGCTGCTGCT 1775
1681 AGTCTCTTTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
1776 AGTCTCTTTCTCTTAACTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
1741 CCTTGA 1746
1836 CCTTGA 1841

RESULT 5
US-10-757-262-45
Sequence 45, Application US/10757262
Publication No. US20040197825A1
GENERAL INFORMATION:
APPLICANT: Karichehli, Venkateswarlu
APPLICANT: Siles-Santiago, Immaculada
APPLICANT: Eliasof, Scott D.

```

/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
/ TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44350, 54181, 211, 5687, 884,
/ TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
/ TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43339, 44373, 51164,
/ TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
/ TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
/ TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
/ TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
/ TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
/ FILE REFERENCE: MP103-007P1R0N0M1N
/ CURRENT APPLICATION NUMBER: US/10/757, 262
/ CURRENT FILING DATE: 2004-01-14
/ PRIOR APPLICATION NUMBER: US 60/440,318
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US 60/444,783
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: US 60/457,901
/ PRIOR FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: US 60/468,775
/ PRIOR FILING DATE: 2003-05-08
/ PRIOR APPLICATION NUMBER: US 60/471,614
/ PRIOR FILING DATE: 2003-05-19
/ PRIOR APPLICATION NUMBER: US 60/478,742
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/488,529
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/491,156
/ PRIOR FILING DATE: 2003-07-30
/ PRIOR APPLICATION NUMBER: US 60/499,594
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 60/506,332
/ PRIOR FILING DATE: 2003-09-26
/ NUMBER OF SEQ ID NOS: 136
/ SOFTWARE: PaedSeq for Windows Version 4.0
/ SEQ ID NO 45
/ LENGTH: 2158
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (96) ... (1841)
/ US-10-757-262-45

Query Match      100.0%; Score 1746; DB 19; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCCAAGGGAATTACTTCATCTGCTTCAACAATGTCCTTTTTCCTGATTCCTCAAGCC 60
DB      96 ATGCCAAGGGAATTACTTCATCTGCTTCAACAATGTCCTTTTTCCTGATTCCTCAAGCC 155
QY      61 CTGTGGGACACACACATGGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAC 120
DB      156 CTGTGGGACACACACATGGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAC 215
QY      121 ACCAGGCTGGATGATTCAGGGCAAGCACTGCTGCTGGGAACCCCTGTGCTGTG 180
DB      216 ACCAGGCTGGATGATTCAGGGCAAGCACTGCTGCTGGGAACCCCTGTGCTGTG 275
QY      181 AACCTGTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTACGAAC 240
DB      276 AACCTGTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTACGAAC 335
QY      241 CCGCAGCTCGATGCGCTGGGATTAATTGCGAAGCCCTTCATCCCTAATTTGTGC 300
DB      336 CCGCAGCTCGATGCGCTGGGATTAATTGCGAAGCCCTTCATCCCTAATTTGTGC 395
QY      301 CTCAGAACTCAGAGTGTCTTTAGATCAACATGCTCAAGGTGATTAACCGGAA 360
DB      396 CTCAGAACTCAGAGTGTCTTTAGATCAACATGCTCAAGGTGATTAACCGGAA 455
QY      361 TTGGAGTGTCAAGACTGCTCTTACCTGAACATCTATGCGCTGCCACCGCATACA 420
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
/ TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44350, 54181, 211, 5687, 884,
/ TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
/ TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43339, 44373, 51164,
/ TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
/ TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
/ TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
/ TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
/ TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
/ FILE REFERENCE: MP103-007P1R0N0M1N
/ CURRENT APPLICATION NUMBER: US/10/757, 262
/ CURRENT FILING DATE: 2004-01-14
/ PRIOR APPLICATION NUMBER: US 60/440,318
/ PRIOR FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: US 60/444,783
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: US 60/457,901
/ PRIOR FILING DATE: 2003-03-27
/ PRIOR APPLICATION NUMBER: US 60/468,775
/ PRIOR FILING DATE: 2003-05-08
/ PRIOR APPLICATION NUMBER: US 60/471,614
/ PRIOR FILING DATE: 2003-05-19
/ PRIOR APPLICATION NUMBER: US 60/478,742
/ PRIOR FILING DATE: 2003-06-16
/ PRIOR APPLICATION NUMBER: US 60/488,529
/ PRIOR FILING DATE: 2003-07-18
/ PRIOR APPLICATION NUMBER: US 60/491,156
/ PRIOR FILING DATE: 2003-07-30
/ PRIOR APPLICATION NUMBER: US 60/499,594
/ PRIOR FILING DATE: 2003-09-02
/ PRIOR APPLICATION NUMBER: US 60/506,332
/ PRIOR FILING DATE: 2003-09-26
/ NUMBER OF SEQ ID NOS: 136
/ SOFTWARE: PaedSeq for Windows Version 4.0
/ SEQ ID NO 45
/ LENGTH: 2158
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (96) ... (1841)
/ US-10-757-262-45

Query Match      100.0%; Score 1746; DB 19; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCCAAGGGAATTACTTCATCTGCTTCAACAATGTCCTTTTTCCTGATTCCTCAAGCC 60
DB      96 ATGCCAAGGGAATTACTTCATCTGCTTCAACAATGTCCTTTTTCCTGATTCCTCAAGCC 155
QY      61 CTGTGGGACACACACATGGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAC 120
DB      156 CTGTGGGACACACACATGGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGGAC 215
QY      121 ACCAGGCTGGATGATTCAGGGCAAGCACTGCTGCTGGGAACCCCTGTGCTGTG 180
DB      216 ACCAGGCTGGATGATTCAGGGCAAGCACTGCTGCTGGGAACCCCTGTGCTGTG 275
QY      181 AACCTGTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTACGAAC 240
DB      276 AACCTGTCTCCGAGTCCCTTTGCTGCTCCCGCTGGGATCCCTGCGATTTACGAAC 335
QY      241 CCGCAGCTCGATGCGCTGGGATTAATTGCGAAGCCCTTCATCCCTAATTTGTGC 300
DB      336 CCGCAGCTCGATGCGCTGGGATTAATTGCGAAGCCCTTCATCCCTAATTTGTGC 395
QY      301 CTCAGAACTCAGAGTGTCTTTAGATCAACATGCTCAAGGTGATTAACCGGAA 360
DB      396 CTCAGAACTCAGAGTGTCTTTAGATCAACATGCTCAAGGTGATTAACCGGAA 455
QY      361 TTGGAGTGTCAAGACTGCTCTTACCTGAACATCTATGCGCTGCCACCGCATACA 420
```

Db	1536	GAAGAGCCAGGAGGAGAGAAAGTTACTGAGCCGGAAGATGATGAATACTAGGGCTAC	1595
Qy	1501	TTTGCTCGAACCCGGGAATCTTAATGGAAAGCATCTGTCTCTGTGGCCAGCTTAATATCTG	1560
Db	1596	TTTGCTCGAACCCGGGAATCTTAATGGAAAGCATCTGTCTCTGTGGCCAGCTTAATATCTG	1655
Qy	1561	ACTGAGCAGTACCTTCAGCTGAGCTTGAAACAATGAGCCTCGGACAGAGACTCAAGAAACCG	1620
Db	1556	ACTGAGCAGTACCTTCAGCTGAGCTTGAAACAATGAGCCTCGGACAGAGACTCAAGAAACCG	1715
Qy	1621	CGGGTGCATTTTTGGACCAAGACCATCCCTGATCCTGTCTGTGCTCCGAATGCTCCAC	1680
Db	1716	CGGGTGCATTTTTGGACCAAGACCATCCCTGATCCTGTGTGCTCCGAATGCTCCAC	1775
Qy	1681	AGTCCTCTTTCCTCTTAATCTTCTCTCTCTCTCCAGACCTTCTTTTCTTTTGTGCT	1740
Db	1776	AGTCCTCTTTCCTCTTAATCTTCTCTCTCTCTCTCCAGACCTTCTTTTCTTTTGTGCT	1835
Qy	1741	CCCTGA	1746
Db	1836	CCCTGA	1841

RESULT 6
US-10-45

```

1      Sequence 40, Application/US/10451168
2      Publication No. US20040091969A1
3
4      GENERAL INFORMATION:
5
6      APPLICANT: SMITHKLINE BEECHAM CORPORATION
7      APPLICANT: SMITHKLINE BEECHAM P.L.C.
8      APPLICANT: GLAXO GROUP LIMITED
9      TITLE OF INVENTION: NOVEL COMPOUNDS
10     FILE REFERENCE: GPE0039
11
12     CURRENT FILING DATE: US/10/451,168
13
14     PRIOR APPLICATION NUMBER: US/2003-11-12
15
16     PRIOR FILING DATE: 2003-11-12
17
18     PRIOR APPLICATION NUMBER: PCT/US01/49232
19
20     PRIOR FILING DATE: 2000-12-17
21
22     PRIOR APPLICATION NUMBER: 60/256,710
23
24     PRIOR FILING DATE: 2000-12-19
25
26     PRIOR APPLICATION NUMBER: 60/257,048
27
28     PRIOR FILING DATE: 2000-12-20
29
30     PRIOR APPLICATION NUMBER: 60/260,482
31
32     PRIOR FILING DATE: 2001-01-09
33
34     PRIOR APPLICATION NUMBER: 60/264,922
35
36     PRIOR FILING DATE: 2001-01-30
37
38     PRIOR APPLICATION NUMBER: 60/266,797
39
40     PRIOR FILING DATE: 2001-02-06
41
42     PRIOR APPLICATION NUMBER: 60/276,988
43
44     PRIOR FILING DATE: 2001-03-19
45
46     PRIOR APPLICATION NUMBER: 60/281,535
47
48     PRIOR FILING DATE: 2001-04-04
49
50     PRIOR APPLICATION NUMBER: 60/289,622
51
52     PRIOR FILING DATE: 2002-06-28
53
54     NUMBER OF SEQ ID NOS: 110
55
56     SOFTWARE: FastSeq for Windows Version 4.0.0
57
58     SEQ ID NO 40
59
60     LENGTH: 1728
61
62     TYPE: DNA
63
64     ORGANISM: Homo sapiens
65
66     OS-10-451-168-40

```

Query Match	94.7%	Score 1652.8	DB 18	Length 1728
Best Local Similarly	99.9%	Pred. No. 0		
Matches 1654; Conservative	0	Mismatches	2	Indels 0
				Gaps 0

Oy 91 GGGCTTTGTGCTGAAGGGGCA CAGAGGAA CACAGGCTGGATGATTCAGGGCAAGCA 150
 Db 73 GGGCTTTGTGCTGAAGGGGCA CAGAGGAA CACAGGCTGGATGATTCAGGGCAAGCA 132
 Oy 151 GTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTATTCCTGGAGTCCCTTTGTGCT 210
 Db 133 GTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTATTCCTGGAGTCCCTTTGTGCT 192

QY	211	CCCCCGCGGGGATCCCTCGGCAATTTAACGAACCCGGAGCCTGATGCGCCCTGGGATTAACCTTG	270
Db	193	CCCCCGCGGGGATCCCTCGGCAATTTAACGAACCCGGAGCCTGATGCGCCCTGGGATTAACCTTG	252
QY	271	CGAGAGGCCACCTCTCAACCTTAATTTGTGCGCTCCAGAACTCAGAGTGCGCTCTCTAGAT	330
Db	253	CGAGAGGCCACCTCTCAACCTTAATTTGTGCGCTCCAGAACTCAGAGTGCGCTCTCTAGAT	312
QY	331	CAACACATGCTCAAGAGTGCAATTACCCGAATTCGAGGTGACAGAAAGACTGCGCTTCACCTG	390
Db	313	CAACATATGCTCAAGAGTGCAATTACCCGAATTCGAGGTGACAGAAAGACTGCGCTTCACCTG	372
QY	391	AAACATCTATGCGGCTCGGCCCAAGCCGATACAGGCTCCCAAGCTCCCCGCTTGTGTGTGCTTC	450
Db	373	AAACATCTATGCGGCTCGGCCCAAGCCGATACAGGCTCCCAAGCTCCCCGCTTGTGTGTGCTTC	432
QY	451	CCAGAGGGTGCCTTCAAGACTGCGCTCAGCGCTCCATCTTTGATGAGGTGCGGCTCGTGCCTGC	510
Db	433	CCAGAGGGTGCCTTCAAGACTGCGCTCAGCGCTCCATCTTTGATGAGGTGCGGCTCGTGCCTGC	492
QY	511	TATGAGGACGTGCTGTGTGTGTGTGCTGCTCAGTACCGGCTCAGGAATATTTGGTTCTTACCC	570
Db	493	TATGAGGACGTGCTGTGTGTGTGTGCTGCTCAGTACCGGCTCAGGAATATTTGGTTCTTACCC	552
QY	571	ACATGGGATCAGCATGCTCCCGGGGACCTGGGCTTCAAGGACCAAGGTGCGCTCTGCTGCC	630
Db	553	ACATGGGATCAGCATGCTCCCGGGGACCTGGGCTTCAAGGACCAAGGTGCGCTCTGCTGCC	612
QY	631	TGGGTCCAGAAAGAACATCGAGTTCTTCGCGTGGGAGACCCAGCTCGTACCATCTTTGGCC	690
Db	613	TGGGTCCAGAAAGAACATCGAGTTCTTCGCGTGGGAGACCCAGCTCTGTGACCATCTTTGGCC	672
QY	691	GAGTCGCGGGAGGCCATPAAGTGTTCATAGTCTTAACTGTCTCCATGCGCCAAAGGCTTAA	750
Db	673	GAGTCGCGGGAGGCCATPAAGTGTTCATAGTCTTAACTGTCTCCATGCGCCAAAGGCTTAA	732
QY	751	TTCCACAAGCCCATCATGAGAGTGCGGGTGGCCATCATCCCTTACCTGAGAGGCCCATGAT	810
Db	733	TTCCACAAGCCCATCATGAGAGTGCGGGTGGCCATCATCCCTTACCTGAGAGGCCCATGAT	792
QY	811	TATAGGAAGAGTGAGGACCTGCAAGTGGTGGCAATTTCTGTGGTAAACAATGGCTCAGAC	870
Db	793	TATAGGAAGAGTGAGGACCTGCAAGTGGTGGCAATTTCTGTGGTAAACAATGGCTCAGAC	852
QY	871	TCTGAGGCCCTCGCTGAGAGTGCGCTGAGGACAAACCCCTCACAAGAGCTGCTGACCCCTGAC	930
Db	853	TCTGAGGCCCTCGCTGAGAGTGCGCTGAGGACAAACCCCTCACAAGAGCTGCTGACCCCTGAC	912
QY	931	CAGAAAACAAAAGTCTTTCATCTGAGTGGTGTGAGTGTGCTTTCTTCTAATGAGCCTCTA	990
Db	913	CAGAAAACAAAAGTCTTTCATCTGAGTGGTGTGAGTGTGCTTTCTTCTAATGAGCCTCTA	972
QY	991	GATCTATTTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCCATCATGCGAGTCAATTAACAC	1050
Db	973	GATCTATTTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCCATCATGCGAGTCAATTAACAC	1032
QY	1051	GAGGTGTGAGCTTCTGCTGCGCTCATGAGAGGAGGCTCCTGAGATCTCTCAGTGGCTCAACAAG	1110
Db	1033	GAGGTGTGAGCTTCTGCTGCGCTCATGAGAGGAGGCTCCTGAGATCTCTCAGTGGCTCAACAAG	1092
QY	1111	TCCCTTGCCTCCATCTGATACAAACATCTGACATCCCGCTCAGTATTTTGGCACCTT	1170
Db	1093	TCCCTTGCCTCCATCTGATACAAACATCTGACATCCCGCTCAGTATTTTGGCACCTT	1152
QY	1171	GTCGCTAATGAATCTTCCATGACAGACACTCCCTGACTGAAATCCGAGACAGTCTTCTG	1230
Db	1153	GTCGCTAATGAATCTTCCATGACAGACACTCCCTGACTGAAATCCGAGACAGTCTTCTG	1212
QY	1231	GACTTGTCTGGAGATGTGTTCTTTGTGTGTCCTGCACTGATCAAGCTCGATATACAGAGA	1290
Db	1213	GACTTGTCTGGAGATGTGTTCTTTGTGTGTCCTGCACTGATCAAGCTCGATATACAGAGA	1272
QY	1291	GATCTGTGTGACCTGTCTACTTCTAATGAGTTTCGAGACCGGCTCAGTGTCTTTGAAGAC	1350

Db 1273 GATGCTGGTGCACTGCTACTCTTATGAGTTTCGGCACCGGCTTCAGTGCCTTGTGAAGAC 1332

Qy 1351 ACGAAGCCGGCTTTTGTCTAAAGCCGACACGCTGATGGAAGTCCGCTTTGTGTGGTGGT 1410

Db 1333 ACGAAGCCAGCCTTTTGTCTAAACCCACACGCTGATGGAAGTCCGCTTTGTGTGGTGGT 1392

Qy 1411 GCCTTCTCGAAGGGGGACATTTGTAATGTTGGAAGGACACGAGAGAGAGATTAATG 1470

Db 1393 GCCTTCTCGAAGGGGGACATTTGTAATGTTGGAAGGACACGAGAGAGAGATTAATG 1452

Qy 1471 AGCGGGAAGATGATGAATTAATCTGGGCTACCTTTGCTCGAACCGGGAAATCTTAATGGAA 1530

Db 1453 AGCGGGAAGATGATGAATTAATCTGGGCTACCTTTGCTCGAACCGGGAAATCTTAATGGAA 1512

Qy 1531 GACCTGTCTCTGTGGCCGACCTTATTAATGTACTGAGCAGTACCTTCAGACTGGACTTGAAC 1590

Db 1513 GACCTGTCTCTGTGGCCGACCTTATTAATGTACTGAGCAGTACCTTCAGACTGGACTTGAAC 1572

Qy 1591 ATGAGCCTCGGACAGAGACTCAAAAGAACCGGGGGGATTTTGTGAAACAGACCATCCCC 1650

Db 1573 ATGAGCCTCGGACAGAGACTCAAAAGAACCGGGGGGATTTTGTGAAACAGACCATCCCC 1632

Qy 1651 CTGATTCCTGTCTGCTCCGACATGCTCAAGTCTCTTTTCTTCTTAACTTCTCTCT 1710

Db 1633 CTGATTCCTGTCTGCTCCGACATGCTCAAGTCTCTCTTTTCTTCTTAACTTCTCTCT 1692

Qy 1711 CTGCTCGAGCCTTTCTTTTCTTTTGTGCTCTTGA 1746

Db 1693 CTGCTCGAGCCTTTCTTTTCTTTTGTGCTCTTGA 1728

```

RESULT 7
US-10-433--256--23
/ Sequence 23, Application US/10433256
/ Publication No. US20040081980A1
/ GENERAL INFORMATION:
/ APPLICANT: SANJANWALA, Madhusudan M.; YAO, Monique G.
/ APPLICANT: AU-YOUNG, Janice K.; BAUGHN, Mariah R.
/ APPLICANT:
/ APPLICANT:
/ APPLICANT:
/ APPLICANT:
/ APPLICANT: LEE, Ernestine A.; DING, Li
/ APPLICANT: ARVIZU, Chandra S.; RING, HuiJun Z.
/ APPLICANT: HAPFALLA, April J.A.; TANG, Y. Tom
/ APPLICANT: YUE, Henry; TRIBOUTLEY, Catherine M.
/ APPLICANT: LU, Dying; ALINA M.; LAL, Preeti G.
/ APPLICANT: WARREN, Bridget A.; YANG, Junming
/ APPLICANT: CHAWLA, Narinder K.; NGUYEN, Damien B.
/ APPLICANT: GANDHI, Ameena R.; LU, Yan
/ APPLICANT: ISON, Craig H.
/ TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
/ FILE REFERENCE: PI-0313 USN
/ CURRENT APPLICATION NUMBER: US/10/433,256
/ CURRENT FILING DATE: 2003-05-30
/ PRIOR APPLICATION NUMBER: PCT/US01/47429
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: US 60/254,308
/ PRIOR FILING DATE: 2000-12-08
/ PRIOR APPLICATION NUMBER: US 60/256,189
/ PRIOR FILING DATE: 2000-12-15
/ PRIOR APPLICATION NUMBER: US 60/257,713
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: US 60/262,706
/ PRIOR FILING DATE: 2001-01-19
/ PRIOR APPLICATION NUMBER: US 60/266,020
/ PRIOR FILING DATE: 2001-02-02
/ NUMBER OF SEQ ID NOS: 26
/ SOFTWARE: PERL Program
/ SEQ ID NO 23
/ LENGTH: 2232
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No: 6538080CB1

```

US-10-433-256-23

Query Match	94.6%	Score 1651.6;	DB 18;	Length 2232;
Best Local Similarity	99.5%	Pred. No. 0;		
Matches 1657; Conservative	0;	Mismatches	9;	Gaps 0

QY	81	GGGAAAAA	CTGGG	CCCTTCTGCTGAAGGG	CCACAGAGAAC	ACCAGGCTGGATGATTTCA	144	
Db	264	GGTCCAGAGAGGG	CCCTTCTGCTGAAGGG	CCACAGAGAA	CACACAGGCTGGATGATTTCA	322		
QY	141	GGGCAAGCAAGTCA	CTGTGTGGAGAAC	CCCTGTGCTGAA	CGTGTTCCTCGAGTCC	200		
Db	324	GGGCAAGCAAGTCA	CTGTGTGGAGAAC	CCCTGTGCTGAA	CGTGTTCCTCGAGTCC	38		
QY	201	CTTTCGCTCT	CCCCCGCTGGGAT	TCCTCTGCATTTACGA	ACCCGACGCTGCATCGCCCTG	26		
Db	384	CTTTCGCTCT	CCCCCGCTGGGAT	TCCTCTGCATTTACGA	ACCCGACGCTGCATCGCCCTG	44		
QY	261	GGATTAAC	CTGCGAGAAC	CCACTCTTACCC	TAAATTGTGCTCCGAA	CTCAGAGTGGCT	32	
Db	444	GGATTAAC	CTGCGAGAAC	CCACTCTTACCC	TAAATTGTGCTCCGAA	CTCAGAGTGGCT	50	
QY	321	GCTCTTGAT	ATCAACAT	TGCTCAAGTGCATTTAC	CCGAAATTCGAGTGTGAGAA	AGCTG	38	
Db	504	GCTCTTGAT	ATCAACAT	TGCTCAAGTGCATTTAC	CCGAAATTCGAGTGTGAGAA	AGCTG	56	
QY	381	CCTCTACCTG	AACATCAT	TGATGCGCTCTCCAC	CGCATACAGGCTC	CAAGCTCCCGCAT	44	
Db	564	CCTCTACCTG	AACATCAT	TGATGCGCTCTCCAC	CGCATACAGGCTC	CAAGCTCCCGCAT	62	
QY	441	GGTGTGTT	CCCCAGAGAGT	GCCTTCAAGACT	GGCTTCACATCTTTGATGGGTCCG		50	
Db	624	GGTGTGTT	CCCCAGAGAGT	GCCTTCAAGACT	GGCTTCACATCTTTGATGGGTCCG		68	
QY	501	CCTGGCTG	CCCTATGAGAG	AGTGTGCTGTTGTGCT	CGCATACCGGCGTAAAGATATTTGG		58	
Db	684	CCTGGCTG	CCCTATGAGAG	AGTGTGCTGTTGTGCT	CGCATACCGGCGTAAAGATATTTGG		74	
QY	561	TTTCTTTCA	CCACAT	TGGGATCAGATGCT	CCGGGGAATG	GGGCTTCAAGAAC	CGAGTGGC	62
Db	744	TTTCTTTCA	CCACAT	TGGGATCAGATGCT	CCGGGGAATG	GGGCTTCAAGAAC	CGAGTGGC	80
QY	621	TGCTGTGT	CTGTGGTCCAGAGAA	CATGAGTCTT	CGTGGGAGCC	CGACGCTGTGAC		68
Db	804	TGCTGTGT	CTGTGGTCCAGAGAA	CATGAGTCTT	CGTGGGAGCC	CGACGCTGTGAC		86
QY	681	CATCTTTGG	CGAGTCCG	CGGGAGCCAT	TAAGTGTTC	TATGCTTATAC	GTCTTCCATG	74
Db	864	CATCTTTGG	CGAGTCCG	CGGGAGCCAT	TAAGTGTTC	TATGCTTATAC	GTCTTCCATG	92
QY	741	CAAAAGCTT	ATTTCCAAAG	GCATCAT	TGAGAGTGGGGTGG	CCATACCTTACCTG		80
Db	924	CAAAAGCTT	ATTTCCAAAG	GCATCAT	TGAGAGTGGGGTGG	CCATACCTTACCTG		98
QY	801	GGGCCAT	GATTTATGAA	AGATGAGAC	CTTGCAGTGTG	GCACATTTCTGTGTAA	CAA	86
Db	984	GGGCCAT	GATTTATGAA	AGATGAGAC	CTTGCAGTGTG	GCACATTTCTGTGTAA	CAA	104
QY	861	TGCGTCAG	ACTCTGAGG	CCCTGCTGAGGTG	CTGAGAACAA	ACCTCCAGAGAGTGT		92
Db	1044	TGCGTCAG	ACTCTGAGG	CCCTGCTGAGGTG	CTGAGAACAA	ACCTCCAGAGAGTGT		110
QY	921	GACCTCT	CAGCCGAA	AAACAAAGTCTT	TCACTCGAGTGTG	ATGATGATGATCTT	CTTCTTA	98
Db	1104	GACCTCT	CAGCCGAA	AAACAAAGTCTT	TCACTCGAGTGTG	ATGATGATGATCTT	CTTCTTA	116
QY	981	TGAGCTCT	AGATCTA	TATGTCTGAGAA	AGATTTAAAGCA	TTTCTTCCATATGAGAGT		104
Db	1164	TGAGCTCT	AGATCTA	TATGTCTGAGAA	AGATTTAAAGCA	TTTCTTCCATATGAGAGT		122
QY	1041	CAATTAACA	CGAGTGTG	CTTCTGCTG	CCCTATGAAGAGG	CTCTGAGATCT	CAGTGG	110
Db	1224	CAATTAACA	CGAGTGTG	CTTCTGCTG	CCCTATGAAGAGG	CTCTGAGATCT	CAGTGG	128

QY 1101 CTCGAAGAGTCCCTTGGCCCTCCATCTGATACAAAACATCTGCAATCCCGCTCAGTA 1160
DB 1284 CTCGAAGAGTCCCTTGGCCCTCCATCTGATACAAAACATCTGCAATCCCGCTCAGTA 1343
QY 1161 TTTCGACCTTGTGGCTAATGAATATCTTCATGACAGACATCCCTGATGAAATCCGAGA 1220
DB 1344 TTTCGACCTTGTGGCTAATGAATATCTTCATGACAGACATCCCTGATGAAATCCGAGA 1403
QY 1221 CAGTCTTCTGACCTTGTGGAGATGTGTCTTGTGTGCTCCCTGACATGACAGCTG 1280
DB 1404 CAGTCTTCTGACCTTGTGGAGATGTGTCTTGTGTGCTCCCTGACATGACAGCTG 1463
QY 1281 ATATCAGAGAGATGCTGAGTCACTGTCTATCTTATGAGTTCCGACCGGCTCAGTG 1340
DB 1464 ATATCAGAGAGATGCTGAGTCACTGTCTATCTTATGAGTTCCGACCGGCTCAGTG 1523
QY 1341 CTTTGAAGACAGAAAGCCGCTTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGT 1400
DB 1524 CTTTGAAGACAGAAAGCCGCTTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGT 1583
QY 1401 GTTCCGTGTGCTCTCTGTAAGGGGGAATTTGTATGTTTGAAGAGCCAGGAGAGA 1460
DB 1584 GTTCCGTGTGCTCTCTGTAAGGGGGAATTTGTATGTTTGAAGAGCCAGGAGAGA 1643
QY 1461 GAACTTACTGAGCCGGAAGATGATGAATATCTGAGCTTGTGCTCGAAGCGGGAATCC 1520
DB 1644 GAACTTACTGAGCCGGAAGATGATGAATATCTGAGCTTGTGCTCGAAGCGGGAATCC 1703
QY 1521 TAATGGAGACGACTGTCTGTGTGGCCAGCTTATATCTGACTGACAGTACTCCAGCT 1580
DB 1704 TAATGGAGACGACTGTGTGTGGCCAGCTTATATCTGACTGACAGTACTCCAGCT 1763
QY 1581 GGAATTGAACATGAGCTCTGGAAGAGACTCAAGAACCGCGGGTGGATTTTGGACAG 1640
DB 1764 GGAATTGAACATGAGCTCTGGAAGAGACTCAAGAACCGCGGGTGGATTTTGGACAG 1823
QY 1641 CACCATCCCCCGATCTGTCTGCTCCGACATGCTCCGACAGTCCCTTCTCTTAC 1700
DB 1824 CACCATCCCCCGATCTGTCTGCTCCGACATGCTCCGACAGTCCCTTCTCTTAC 1883
QY 1701 TTTCCTCTCTCTCTCAAGCCCTTCTTCTTGTGTGCTCTTGA 1746
DB 1884 TTTCCTCTCTCTCTCTCAAGCCCTTCTTCTTGTGTGCTCTTGA 1929

RESULT 8
US-10-104-047-249
; Sequence 249, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 249
; LENGTH: 2092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-249

Query Match 93.4%; Score 1631.2; DB 17; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 91 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCAA 150
DB 138 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCAA 197

QY 151 GTCACTGTGTGGGAAGCCCTGTGCTGAAAGTGTCTCGAGTCCCTTGTGCTGCT 210
DB 198 GTCACTGTGTGGGAAGCCCTGTGCTGAAAGTGTCTCGAGTCCCTTGTGCTGCT 257
QY 211 CCCCCGTGGGATCCCTGGATTTACGAACCCGACGCTGATCCGCTGGGATTAATTG 270
DB 258 CCCCCGTGGGATCCCTGGATTTACGAACCCGACGCTGATCCGCTGGGATTAATTG 317
QY 271 CGAAGAACCACTCCCTACCTTAATT-----GTGGCTCCGAATCCAGAGTGGCT 320
DB 318 CGAAGAACCACTCCCTACCTTAATTGTAAGCAAGGTGCTCCAGAACTCAGAGTGGCT 377
QY 321 GCTTTAGATCAACATGCTCAAGTGTGATTAACCGGAATTCGAGTGTCAAGAACTG 380
DB 378 GCTTTAGATCAACATGCTCAAGTGTGATTAACCGGAATTCGAGTGTCAAGAACTG 437
QY 381 CCTTACTGAACATCTATGCGCTGCCCAAGCCGATACAGGCTCCAGCTCCGCTTT 440
DB 438 CCTTACTGAACATCTATGCGCTGCCCAAGCCGATACAGGCTCCAGCTCCGCTTT 497
QY 441 GGTGTGTTCCTCAAGAGGTGCTTCAAGATGGCTCAGCTTCATCTTGAATGGTCCG 500
DB 498 GGTGTGTTCCTCAAGAGGTGCTTCAAGATGGCTCAGCTTCATCTTGAATGGTCCG 557
QY 501 CCTGGCTGCTATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
DB 558 CCTGGCTGCTATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 617
QY 561 TTTCCTTCAACATAGGAGATCAAGATGCTCCGAGGAATGAGGCTTCAAGAACAGGTG 620
DB 618 TTTCCTTCAACATAGGAGATCAAGATGCTCCGAGGAATGAGGCTTCAAGAACAGGTG 677
QY 621 TGTCTGTCTGTGGTTCAGAAAGACATGAGTCTTCCGTGGGAGCCCAAGCTGTGAC 680
DB 678 TGTCTGTCTGTGGTTCAGAAAGACATGAGTCTTCCGTGGGAGCCCAAGCTGTGAC 737
QY 681 CATCTTTGGCAGTCCGCGGAGCATTAAAGTGTCTTCAAGTGTCTTCAAGTGTG 740
DB 738 CATCTTTGGCAGTCCGCGGAGCATTAAAGTGTCTTCAAGTGTCTTCAAGTGTG 797
QY 741 CAAAGCTTATTTCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACTCTGGA 800
DB 798 CAAAGCTTATTTCAAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACTCTGGA 857
QY 801 GGCCCATGATTTATGAAGAGTGAAGACCTGAGGTGTGCAATTTGTGTGTAACA 860
DB 858 GGCCCATGATTTATGAAGAGTGAAGACCTGAGGTGTGCAATTTGTGTGTAACA 917
QY 861 TGGGTCAAGCTGTGAGGCTGTGAGGTGCTGAGAGCAAAACCTCCAGAGGTGCT 920
DB 918 TGGGTCAAGCTGTGAGGCTGTGAGGTGCTGAGAGCAAAACCTCCAGAGGTGCT 977
QY 921 GACCTCAGCCAGAAAAGAAAGTCTTCACTGAGTGTGTAAGTGTCTTCTTCTTA 980
DB 978 GACCTCAGCCAGAAAAGAAAGTCTTCACTGAGTGTGTAAGTGTCTTCTTCTTA 1037
QY 981 TGAAGCTCTAGATCTATGTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAG 1040
DB 1038 TGAAGCTCTAGATCTATGTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAG 1097
QY 1041 CAATTAACAGAGAGTGTGCTTCTGCTGCTGATGAAGAGGCTCCGAGATCTCAGTGG 1100
DB 1098 CAATTAACAGAGAGTGTGCTTCTGCTGCTGATGAAGAGGCTCCGAGATCTCAGTGG 1157
QY 1101 CTCGAAGAGTCCCTTGGCCCTCCATCTGATACAAAACATCTGCAATCCCGCTCAGTA 1160
DB 1158 CTCGAAGAGTCCCTTGGCCCTCCATCTGATACAAAACATCTGCAATCCCGCTCAGTA 1217
QY 1161 TTTCGACCTTGTGGCTAATGAATATCTTCATGACAGACATCTCCCTGATGAAATCCGAGA 1220
DB 1218 TTTCGACCTTGTGGCTAATGAATATCTTCATGACAGACATCTCCCTGATGAAATCCGAGA 1277

```

; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 195
; TYPE: DNA
; LENGTH: 1746
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1744)
; US-10-114-270-195

```

Query Match	Best Local Similarity	92.2%;	Score 1435.6;	DB 18;	Length 1746;
Matches 1466;	Conservative	0;	Pred. No. 0;	Mismatches 59;	Indels 15;
				Gaps	1;
QY	81	GGGAAAAAACTGGGCGCTTCTGTGTCGAAAGGGCCACAGAGAAACCAAGGCTGGGATGGATTCA	140		
DB	69	GGTTCAGAGAGGGCGCTTGTGTGTAAGGGCCACAGAGAAACCAAGCTGGGATGGATTCA	128		
QY	141	GGGCAAGCAAGTCACTGTGTGTGGGAAGCCCTGTGTGTAACGTGTTCCTCGAGTCCC	200		
DB	129	GGGCAAGCAAGTCACTGTGTGTGGGAAGCCCTGTGTGTAACGTGTTCCTCGAGTCCC	188		
QY	201	CTTTGCTCTCCCCGCGTGGGATCCCTGGCCATTACGAACCCGACGCTGATGCCCTG	260		
DB	189	CTTTGCTCTCCCCGCGTGGGATCCCTGGCCATTACGAACCCGACGCTGATGCCCTG	248		
QY	261	GGATTAATCTTGGAGAAAGCCACTCTCAACCCAAATTTGTGCTCCAGAACTCAGAGTGGCT	320		
DB	249	GGATTAATCTTGGAGAAAGCCACTCTCAACCCAAATTTGTGCTCCAGAACTCAGAGTGGCT	308		
QY	321	GCTCTTAGATTAACAATGCTCAAGGTGCATTAACCGAAATTTGGAGTGCAGAGACTG	380		
DB	309	GCTCTTAGATTAACAATGCTCAAGGTGCATTAACCGAAATTTGGAGTGCAGAGACTG	368		
QY	381	CCTCTACCTGAACATCTAATGGGCGTGGCCAGCGCGATACAGGCTTCCACGCTCTTT	440		
DB	369	CCTCTACCTGAACATCTAATGGGCGTGGCCAGCGCGATACAGGCTTCCACGCTCTTT	428		
QY	441	GGTGTGGTTCCAGAGAGTGCCTTTCAAGACTGTGCTACGCTCCATCTTTGATGGGTCCG	500		
DB	429	GGTGTGGTTCCAGAGAGTGCCTTTCAAGACTGTGCTACGCTCCATCTTTGATGGGTCCG	488		
QY	501	CTGGGCTGCTAATGAGGACGTGCTGGTTGTGTGTCACAGTACCGGCTAGAAATATTTGG	560		
DB	489	CTGGGCTGCTAATGAGGACGTGCTGGTTGTGTGTCACAGTACCGGCTAGAAATATTTGG	548		
QY	561	TTTCTTCAACCAATGGGATCAGATGCTCCGGGAACTGGGCTTTAAGAACCAAGTGGC	620		
DB	549	TTTCTTCAACCAATGGGATCAGATGCTCCGGGAACTGGGCTTTAAGAACCAAGTGGC	608		

621 TGCTGTCCTGGGTCCAGAAACATGAGTCTTCCGTTGGGAGACCCAGCTCTGTAC 680
609 TGCTGTCCTGGGTCCAGAAACATGAGTCTTCCGTTGGGAGACCCAGCTCTGTAC 668
681 CATCTTTGGCGAGTCCGGGGAGCCATAAGTGTCTTCTTATCTTATCTGTCTCCATGGC 740
669 CATCTTTGGCGAGTCCGGGGAGCCATAAGTGTCTTCTTATCTTATCTGTCTCCATGGC 728
741 CAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACTGTGA 800
729 CAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATCATCTTACTGTGA 788
801 GGGCCATGTTATGAGAGAGTGAAGACCTTGAGGAGTGTGACATTTCTGTGTAAACA 860
789 GGGCCATGTTATGAGAGAGTGAAGACCTTGAGGAGTGTGACATTTCTGTGTAAACA 848
861 TGCGTCAGACTCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCCTCCAGAGAGTGTCT 920
849 TGCGTCAGACTCTGAGGCGCTGCTGAGGTGCTGAGAGCAAAACCCTCCAGAGAGTGTCT 908
921 GACCTTCAGCCGAAACAAAGTCTTTCACTGAGTGGTTGATGTGTCTTTCTTCTTAA 980
909 GACCTTCAGCCGAAACAAAGTCTTTCACTGAGTGGTTGATGTGTCTTTCTTCTTAA 968
981 TGAGGCTCTAGATCTATGTCTCAGAAAGCATTTAAGCAATTCCTTCATGATGGAGT 1040
969 TGAGGCTCTAGATCTATGTCTCAGAAAGCATTTAAGCAATTCCTTCATGATGGAGT 1028
1041 CAATTAACACAGAGTGGCTTCTGCTGCTTATGAGAGAGGCTCCTGAGATCCTCAGTGG 1100
1029 CAATTAACACAGAGTGGCTTCTGCTGCTTATGAGAGATTCCTGCTGCTCATACTGC 1088
1101 CTCGAAAGTCTCTTGGCTCCATCTGTATCAAA-----ACATCTGCA 1145
1089 CACTCCCTCAAAACCGTGAATGACAGCTTTGGCTTCAACAGCTGGGCAATTCACAGAAAGCA 1148
1146 CATCCCGCTCAGTATTTTGGCACTTGTGGCTAATAATCTTCCATGAGAAACACCTCCCT 1205
1149 TATCCCGCTCAGTATTTTGGCACTTGTGGCTAATAATCTTCCATGAGAAACACCTCCCT 1208
1206 GACTGAAATCCGAGACAGTCTTCTGGAATCTGTGGAAGATGTCTTGTGTGGTCCCTGC 1265
1209 GACTGAAATCCGAGACAGTCTTCTGGAATCTGTGGAAGATGTCTTGTGTGGTCCCTGC 1268
1266 ACTGATCAACAGTCTGATATCAAGAGATGCTGTGACCTGTCTACTTCTATGAGTTTGC 1325
1269 ACTGATCAACAGTCTGATATCAAGAGATGCTGTGACCTGTCTACTTCTATGAGTTTGC 1328
1326 GACCGGCTCAGTCTTGTGAAGACAGAGCCGCTTTTGTCAAAAGCCGACACGCTGA 1385
1339 GACCGGCTCAGTCTTGTGAAGACAGAGCCGCTTTTGTCAAAAGCCGACACGCTGA 1388
1386 TGAAGTCGCTTGTGTGCTGCTGCTTCTTGAAGGGGGGACATTTGTTATGTTGGAAG 1445
1389 TGAAGTCGCTTGTGTGCTGCTGCTTCTTGAAGGGGGGACATTTGTTATGTTGGAAG 1448
1446 AGCCACGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATACTGGGCTACCTTTGC 1505
1449 AGCCACGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATACTGGGCTACCTTTGC 1508
1506 TCGAAGCCGGGAATCTTAATGGGAAGCACTGTCTGTGGCCAGCTTAATACTGACTGA 1565
1509 TCGAAGCCGGGAATCTTAATGGGAAGCACTGTCTGTGGCCAGCTTAATACTGACTGA 1568
1566 GCAATACCTCAGAGTGAATGAAATGAGCTTGGACAGAACTCAAGAAACCGGGAG 1625
1569 GCAATACCTCAGAGTGAATGAAATGAGCTTGGACAGAACTCAAGAAACCGGGAG 1628
1626 GGAATTTTGG 1635
1629 AGATGTGTGG 1638

RESULT 10
US-10-451-168-41
; Sequence 41, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50039
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-451-168-41
Query Match 82.0%; Score 1432.4; DB 18; Length 1746;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 51; Indels 21; Gaps 1;
91 GGGCCCTTCTGCTGAGAGGGGCAAGAGGAAACACAGGCTGGATGATTCAGGGCAAGCA 150
73 GGGCCCTTCTGCTGAGAGGGGCAAGAGGAAACACAGGCTGGATGATTCAGGGCAAGCA 132
151 GTCACTGTCGAGAGGAGCCCTGCTGGAACGTGTTCTCGAGTCCCTTTGCTGCT 210
133 GTCACTGTCGAGAGGAGCCCTGCTGGAACGTGTTCTCGAGTCCCTTTGCTGCT 192
211 CCCCCGTGAGATCCCTGAGATTTAAGAACCGGACGCTGATCGCCCTGGAGTAACTTG 270
193 CCCCCGTGAGATCCCTGAGATTTAAGAACCGGACGCTGATCGCCCTGGAGTAACTTG 252
271 CGAAGAGCACTCTTACCTTAATTTGGCTCCAGAACTGAGAGTGGCTCTTAGAT 330
253 CGAAGAGCACTCTTACCTTAATTTGGCTCCAGAACTGAGAGTGGCTCTTAGAT 312
331 CAACACATGCTCAAGATGATTAACCCGAAATTCGAGAGTCCAGAAAGTCCCTTACCTG 390
313 CAACACATGCTCAAGATGATTAACCCGAAATTCGAGAGTCCAGAAAGTCCCTTACCTG 372
391 AACATCTATGCGCTGCGCCACGCGAGTACAGGCTCCAAAGCTCCCGCTTGTGTGTTTC 450
432 AACATCTATGCGCTGCGCCACGCGAGTACAGGCTCCAAAGCTCCCGCTTGTGTGTTTC 432
451 CGAAGAGTGGCTTCAAGATGAGTCCAGCTCATCTTGTAGTGGTCCGCTGCTGCTGC 510
433 CGAAGAGTGGCTTCAAGATGAGTCCAGCTCATCTTGTAGTGGTCCGCTGCTGCTGC 492
511 TATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570

Db 493 TATGAGAGCTGTGCTGTGTGATCGTCCAGTACCGGCTAGGAATATTGTTCTTACC 552
Qy 571 ACATGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGTCTGCC 630
Db 553 ACATGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGTCTGCC 612
Qy 631 TGGGTCCAGAAACATCGAGTCTTCTGTTGGGAGACCAGCTGTGACCATTTTGGC 690
Db 613 TGGGTCCAGAAACATCGAGTCTTCTGTTGGGAGACCAGCTGTGACCATTTTGGC 672
Qy 691 GAGTCCCGGGAGGACATTAAGTGTTCATGCTTATTAAGTCTGCTCCATGSCAAAGCTTA 750
Db 673 GAGTCCCGGGAGGACATTAAGTGTTCATGCTTATTAAGTCTGCTCCATGSCAAAGCTTA 732
Qy 751 TTCCACAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTAACCTGGAGGCCCATGAT 810
Db 733 TTCCACAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTAACCTGGAGGCCCATGAT 792
Qy 811 TATGAGAAAGTGAAGACCTGCAAGTGTTCACATTTCTGTGTAACAATGCGTCAGAC 870
Db 793 TATGAGAAAGTGAAGACCTGCAAGTGTTCACATTTCTGTGTAACAATGCGTCAGAC 852
Qy 871 TCTGAGGCCCCGTGAGGTGGCCCTGAGAGCAAAACCTCCAAAGAGCTGCTGACCTGACG 930
Db 853 TCTGAGGCCCCGTGAGGTGGCCCTGAGAGCAAAACCTCCAAAGAGCTGCTGACCTGACG 912
Qy 931 CAGAAAACAAAGTCTTCACTCGAGTGTGTGATGTGCTTCTTCTTAATAGCCCTTA 990
Db 913 CAGAAAACAAAGTCTTCACTCGAGTGTGTGATGTGCTTCTTCTTAATAGCCCTTA 972
Qy 991 GATCTATGTGCTCAGAAAGCATTTAAAGCATTCCTTCATCATCGAGTCAATAACAC 1050
Db 973 GATCTATGTGCTCAGAAAGCATTTAAAGCATTCCTTCATCATCGAGTCAATAACAC 1032
Qy 1051 GAGTGTGCTTCTGCTGCTTGAAGAGAGGCTCCGAAATCCTCAGTGGCTTCAACAG 1110
Db 1033 GAGTGTGCTTCTGCTGCTTGAAGAGAGGCTCCGAAATCCTCAGTGGCTTCAACAG 1092
Qy 1111 TCCCTT-----GCCCTCATCTGTATCAAAACATCTCGACATC 1149
Db 1093 AACCGTATGATGACCTTGGCTTCAACAGCTGGGCAATTCACAGAAACATAGACATC 1152
Qy 1150 CCGGCTCAGTATTTGACCTTGGCTTGAATGAATCTTCCATGACAAAGCATCTCTGACT 1209
Db 1153 CCGGCTCAGTATTTGACCTTGGCTTGAATGAATCTTCCATGACAAACATCTCTGACT 1212
Qy 1210 GAAATCCGAGACATCTTCTGACCTTGGAGATGTGTTCTTGTGATCCCTGACATG 1269
Db 1213 GAAATCCGAGACATCTTCTGACCTTGGAGATGTGTTCTTGTGATCCCTGACATG 1272
Qy 1270 ATTCACAGCTCATATCAGAGAGATCTGTGACCTGTCTAATCTTATGATTTCCGAC 1329
Db 1273 ATTCACAGCTCATATCAGAGAGATCTGTGACCTGTCTAATCTTATGATTTCCGAC 1332
Qy 1330 CCGGCTCAGTCTTGAAGACAGAACCGGCTTTGTCAAGCGCAACCAAGCTGATGAA 1389
Db 1333 CCGGCTCAGTCTTGAAGACAGAACCGGCTTTGTCAAGCGCAACCAAGCTGATGAA 1392
Qy 1390 GTCCGCTTGTGTGCTGTGCTTCTTGAAGGGGACATTTGATTTGATTTGAAGAGCC 1449
Db 1393 GTCCGCTTGTGTGCTGTGCTTCTTGAAGGGGACATTTGATTTGATTTGAAGAGCC 1452
Qy 1450 ACCGAGAGAGAAAGTTACTGAGCGGAGAGATGATGAATATCTGGGCTACCTTTGCTGA 1509
Db 1453 ACCGAGAGAGAAAGTTACTGAGCGGAGAGATGATGAATATCTGGGCTACCTTTGCTGA 1512
Qy 1510 ACCGAGATCTTAATGAGAAAGCACTGTCTGTGCGACAGCTTAATCTGACAGAGCAG 1569
Db 1513 ACCGAGATCTTAATGAGAAAGCACTGTCTGTGCGACAGCTTAATCTGACAGAGCAG 1572
Qy 1570 TACTTCAAGCTGACATTTGAACATGAGCTTGGACAGAGACTCAAAAGAACCGGAGTGGAT 1629
Db 1573 TACTTCAAGCTGACATTTGAACATGAGCTTGGACAGAGACTCAAAAGAACCGGAGAGAT 1632

Qy 1630 TTTTGG 1635
Db 1633 GTGTGG 1638

RESULT 11
US-10-381-898-20
; Sequence 20, Application US/10381898
; Publication No. US2004008687A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valida; BAUGHN, Mariah R.;
; APPLICANT: BOROMSKI, Mark L.; DING, Li;
; APPLICANT: DUGGAN, Brendan; ELLIOTT, Vicki S.;
; APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KHAN, Farrah A.; LAL, Preeti G.;
; APPLICANT: LEE, Ernestine A.; LU, Dyrung Alina M.;
; APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
; APPLICANT: RAMKUMAR, Jayalaxmi; KING, Huijun Z.;
; APPLICANT: SANJANMALA, Madhusudan M.;
; APPLICANT: TANG, Y. Tom; TRIBOUNLEY, Catherine M.;
; APPLICANT: CHAWLA, Nandinder K.; WALSH, Roderick T.;
; APPLICANT: WARREN, Bridget; XU, Yuming; YUE, Henry
; APPLICANT: YANG, Junning; YAO, Monique; YUE, Henry
; TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
; FILE REFERENCE: PI-0233 USN
; CURRENT APPLICATION NUMBER: US/10/381,898
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: PCT/US01/30662
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,947
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,864
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/242,323
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/247,581
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,519
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,834
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,567
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7473645CB1
US-10-381-898-20

Query Match 79.4%; Score 1386.2; DB 18; Length 1857;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

Qy 81 GGGAAAACTGGGCTTCTGTGTAAGGGCCACAGAGAAACACAGCTGGAGTGTCA 140
Db 264 GGTCCAAAGAGGGCTTCTGTGTAAGGGCCACAGAGAAACACAGCTGGAGTGTCA 323
Qy 141 GGGCAAAGCAATGCTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTTCTTGGAGTCC 200
Db 324 GGGCAAAGCAATGCTGTGCTGGAAAGCCCTGTGCTGTGAAAGTGTCTTGGAGTCC 383
Qy 201 CTTTGTGCTCCCGCTGGAGATCCCTGAGATTTACGAACCGGAGCTGATGAGCCCTG 260
Db 384 CTTTGTGCTCCCGCTGGAGATCCCTGAGATTTACGAACCGGAGCTGATGAGCCCTG 443
Qy 261 GGATTAATTGGAGAGCACTCTTACCTTAATTGTGCTCCAGAACTCAGAGTGGCT 320

```
Db 444 GATTAATCTGGAGAGAGCCACTCTACCTTAATTTGGCCCTCCAGAACTCAGAGTGGCT 503
Qy 321 GCTCTTAATCAACACATGCTCAAGAGTGAATTAACCCGAATTCGAGAGTCAAGAGACTG 380
Db 504 GCTCTTAATCAACACATGCTCAAGAGTGAATTAACCCGAATTCGAGAGTCAAGAGACTG 563
Qy 381 CCTCACTCAAGCAATCTATGAGCTGAGCCGATCAAGAGTCAAGAGTCAAGAGTCAAGAGT 440
Db 564 CCTCACTCAAGCAATCTATGAGCTGAGCCGATCAAGAGTCAAGAGTCAAGAGTCAAGAGT 623
Qy 441 GATGAGTCTCCAGAGAGTGGCTTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGT 500
Db 624 GATGAGTCTCCAGAGAGTGGCTTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGT 683
Qy 501 CTTGGCTGCTCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 560
Db 684 CTTGGCTGCTCAATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 743
Qy 561 TTTCTTCAACCAATGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAG 620
Db 744 TTTCTTCAACCAATGAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAG 803
Qy 621 TGGCTGCTGCTGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 680
Db 804 TGGCTGCTGCTGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCA 863
Qy 681 CATCTTTGCGAGTCCGCGGAGCCATTAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 740
Db 864 CATCTTTGCGAGTCCGCGGAGCCATTAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 923
Qy 741 CAAAGGCTTATTCACAAAGCCATCAATGAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG 800
Db 924 CAAAGGCTTATTCACAAAGCCATCAATGAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG 983
Qy 801 GGGCCATGATTAATGAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 860
Db 984 GGGCCATGATTAATGAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 1043
Qy 861 TGGCTGAGACTGAGAGCTGCTGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 920
Db 1044 TGGCTGAGACTGAGAGCTGCTGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 1103
Qy 921 GACCTCGAGCAGAGAGAGAGTCTTCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 980
Db 1104 GACCTCGAGCAGAGAGAGAGTCTTCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1163
Qy 981 TGAGGCTCTAGATCTATGCTCAAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 1040
Db 1164 TGAGGCTCTAGATCTATGCTCAAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 1223
Qy 1041 CAATTAACAGAGAGTGGCTTCTGCTGCTGAAGAGTCTGAGATCTCTGAGATCTCTGAG 1100
Db 1224 CAATTAACAGAGAGTGGCTTCTGCTGCTGA----- 1255
Qy 1101 CTCGCAAGAGTCCCTTGGCCCTTCATCTGATCAAAACATCTCGACATCCCGCCCTCAGTA 1160
Db 1256 -----TGCAATCCCGCCCTCAGTA 1274
Qy 1161 TTTGCACTTGGGCTAATGAATATCTTCATGACAGAGCTCCGAGCTGAATCCGAGA 1220
Db 1275 TTTGCACTTGGGCTAATGAATATCTTCATGACAGAGCTCCGAGCTGAATCCGAGA 1334
Qy 1221 CAGTCTTCTGAGCTTCTGAGAGTGTCTTTGTGTGTCTCTGAGCTGATCAAGCTCG 1280
Db 1335 CAGTCTTCTGAGCTTCTGAGAGTGTCTTTGTGTGTCTCTGAGCTGATCAAGCTCG 1394
Qy 1281 AATATCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1340
Db 1395 AATATCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1454
Qy 1341 CTTTGAAGACAGAGCGGCTTTTGTCAAGCCGACAGCTGATGAGTCCGCTTTGT 1400
```

```
Db 1455 CTTTGAAGACAGAGCGGCTTTTGTCAAGCCGACAGCTGATGAGTCCGCTTTGT 1514
Qy 1401 GTTGGGTGTGCTCTTCTGAAGAGGAGCATTTGTTATGTTCAAGAGGCAAGAGAGGA 1460
Db 1515 GTTGGGTGTGCTCTTCTGAAGAGGAGCATTTGTTATGTTCAAGAGGCAAGAGAGGA 1574
Qy 1461 GAAGTTACTGAGCCGAGAGATGATGAAATTAAGTGGCTACTTTTGTCTGAAACGGGAATCC 1520
Db 1575 GAAGTTACTGAGCCGAGAGATGATGAAATTAAGTGGCTACTTTTGTCTGAAACGGGAATCC 1634
Qy 1521 TAATGGAAACAGCTGTCTCTGTGGCCAGCTTATTAATCTGACTGAGCACTTCAAGT 1580
Db 1635 TAATGGAAACAGCTGTCTCTGTGGCCAGCTTATTAATCTGACTGAGCACTTCAAGT 1694
Qy 1581 GACTTGAACATGAGCCCTCGAGCAGAGACTCAAGAGACCGGGGGTGAATTTTGG 1635
Db 1695 GACTTGAACATGAGCCCTCGAGCAGAGACTCAAGAGACCGGGAGAGATGTGTGG 1749

RESULT 12
US-10-094-749-736
; Sequence 736, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKI, ICHIRO
; APPLICANT: SEKI, NAOHKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOMYUKI
; APPLICANT: NAGAHARI, KENDI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 736
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-736

Query Match 77.0%; Score 1344.4; DB 17; Length 1962;
Beet Local Similarity 90.9%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

Qy 91 GGGCTTTTGTCTGAAGAGCCACAGAGAAACACAGAGTGGATGATTCAGAGGCAAGCA 150
Db 222 GGGCTTTTGTCTGAAGAGCCACAGAGAAACACAGAGTGGATGATTCAGAGGCAAGCA 281
Qy 151 GTCACTGTCTGAGAGAGCCCTGTGCTGTGAACGTGTTCTTCTGAGAGTCCCTTTTGTCT 210
Db 282 GTCACTGTCTGAGAGAGCCCTGTGCTGTGAACGTGTTCTTCTGAGAGTCCCTTTTGTCT 341
Qy 211 CCCCCGCTGGAGATCCCTGCGATTTACGAAACCGAGCTGATGAGCTCCCTGGAGATACTTG 270
Db 342 CCCCCGCTGGAGATCCCTGCGATTTACGAAACCGAGCTGATGAGCTCCCTGGAGATACTTG 401
```


QY 271 CGAAGACCACTCTCACTCAATTTGTCCTCAGAACTCAGAGTGTGCTTTAGAT 330
 DB 418 CGAAATGCCACATCTTAACCTTAAATTATGCTTCAGAGCTTGAAGTGGCTTCCTAT 477
 QY 331 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAACTGCTTACCTG 390
 DB 478 CAACACGTTCTCAAGGTGCTTACCGCAATTTGAAGCTCCGAAAGACTGCTGTATCTT 537
 QY 331 AACATCTATGCGCTGCGCCAGCCGATACAGAGCTCCAGCTCCCGTCTTGGTGTTC 450
 DB 538 AACATCTATGCGCTGCGCCAGCCGATACAGAGCTCCAGCTCCCGTCTTGGTGTTC 597
 QY 451 CCAGAGGTGCTTCAAGACTGCTCAGCTCCTCATCTTGTATGGGTCCGCTGCTGCTC 510
 DB 598 CCGGGGGTGTCTTCAAGATGGGCTTCAAGTCTCTCTCTTGAATGGTCCGCTTGGCTGC 657
 QY 511 TATGAGACGTGCTGT 570
 DB 658 TACGAGACGTGCTGT 717
 QY 571 ACATGGGATCAGCATGCTCCGCGGAACTGGGCTTCAAGAACCAAGTGTGCTGTCTC 630
 DB 718 ACAGGGATGAGCATGCTCCGCGGAACTGGGCTTGTGGACCAAGTGTGCTCCCTAAC 777
 QY 631 TGGGTCCAGAAAGCATGAGTCTTCCGTTGGGGAACCCAGCTCTGTGACCATCTTGGC 690
 DB 778 TGGGTCCGCGGACAAATGAGTCTTCCGTTGGGGAACCCAGCTCTGTGACCATCTTGGC 837
 QY 691 GAGTCCGCGGAGCCATAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 750
 DB 838 GAGTCCGCGGAGCCATAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 897
 QY 751 TTCCCAAAAGCATGAGAGTGGGTGGCTTCAATCTCTTACCTGAG--GCCAT 807
 DB 898 TTCCCAAAAGCATGAGAGTGGGTGGCTTCAATCTCTTACCTGAGAGCCCTT 957
 QY 808 GATTATGAGAGTGAAGAGCCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 867
 DB 958 GGTGATGAGAGTGAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1017
 QY 868 GACTGTGAGGCTCTGAGT 927
 DB 1018 GACTGTGAGGCTCTGAGT 1077
 QY 928 AGCCGAAAGAAAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 987
 DB 1078 AGCAAGAAAGTCTTGT 1137
 QY 988 CTAGATCTATTGTCTCAGAAAGATTAAGCAATTCCTTCATCATCGGAGTCAATTA 1047
 DB 1138 GTAGCCCTATTGTCTCAGAAAGATTAAGCAATTCCTTCATCATCGGAGTCAATTA 1197
 QY 1048 CACGAGTGTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1107
 DB 1198 CACGAGTGTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1254
 QY 1108 AAGTCTGTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1167
 DB 1255 AAGTCTGTGCTTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1314
 QY 1168 CTGT 1227
 DB 1315 CTGT 1374
 QY 1228 CTGT 1287
 DB 1375 CTGT 1434
 QY 1288 AAGATGT 1347
 DB 1435 AAGATGT 1494
 QY 1348 GACAGAAAGCGGCTTTGTCAAAGCCGACAGGCTGATGAAGTGTGTGTGTGTGTGT 1407

DB 1495 GACAGAGGCGAGCTTTGCTGAAGAACCGATCACTGTGAATGAATCGCTTGTGTGA 1554
 QY 1408 GGTGCTTCTCTAAGGGGACATTTGTATTTTCGAAGAGCCACGAGAGAGAAATTA 1467
 DB 1555 GGTGCTTCTCTAAGGGGACATTTGTATTTTCGAAGAGCCACGAGAGAGAAATTA 1614
 QY 1468 CTGAGCCGGAAGATGATGAATACTGAGCTACCTTTGTCTGACACCGGAAATCTTA 1527
 DB 1615 CTGAGCAGAGAGATGATGAATACTGAGCTACCTTTGTCTGACACCGGAAATCTTA 1674
 QY 1528 AACGACTGTCTCTGTGCGACCTTATATCTGACTGACAGTACCTTCAAGTGTG 1587
 DB 1675 GAAAGTGTGCTCTGTGCGACCTTACACCGAGAGAGCAATCTGAAGCTGATTTG 1734
 QY 1588 AACATGAGCTTGTGACAGAGATCTAAAGAACCGCGGTGTGATTTTGTGACAGCA 1646
 DB 1735 AGTGTGAGGTGTGTGACAGAACTGAAGAGCAAGAGTGTGTGATGAATACAT 1794
 QY 1647 -CCCCGTGATCTGTGTGCTGCGACATGCTCAGATGCTCTTCTTCTTAACCTTCC 1705
 DB 1795 GTCCCTGATACCCCTCACCTTCAGAGGCTTCCAGTCTCTTCTTCTTCTCTCC 1854
 QY 1706 TCTCTCTCTCCAGCTTTCTTTTCTTTTGTGTCTCTTG 1745
 DB 1855 TTCTTGTCTCCGCTGCTGCTTCTTCTTCTGCTCATG 1894

RESULT 14
 US-10-233-933A-1
 ; Sequence 1, Application US/10233933A
 ; Publication No. US2004021417A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamashita, Teisuro
 ; APPLICANT: Miyazaki, Masao
 ; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER
 ; FILE REFERENCE: SHIG FP02US006
 ; CURRENT APPLICATION NUMBER: US/10/233, 933A
 ; PRIOR FILING DATE: 2002-09-03
 ; PRIOR APPLICATION NUMBER: JP2002-057908
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1629
 ; TYPE: DNA
 ; ORGANISM: Felis catus
 US-10-233-933A-1

Query Match 60.9%; Score 1063.2; DB 20; Length 1629;
 Best Local Similarity 80.9%; Pred. No. 0;
 Matches 165; Conservative 0; Mismatches 293; Indels 6; Gaps 2;
 QY 91 GGGCTTCTGTGAGAGGCGACAGAGAAACACAGAGTGTGATTCAGGGCAAGCA 150
 DB 64 GGGCGAGCTGTGATGACACAGTGAAGACACAGGCTGTGATGGTCCGGGGAAGCA 123
 QY 151 GTACTGTGTGGAAGCCGTGTGTGAAGCTGTTCCTGGAAGTCCCTTGTGCT 210
 DB 124 ACCACTGTACTGGAAGACCGTGTGTGAACATGTTCCTGTGGATCCCTATGTGCA 183
 QY 211 CCCCCGTGTGATCCCTGCGATTTACGAACCCGAGCCGTGATTCGCCGTGGATTA 270
 DB 184 CTTCTCTAAGGCCCCCTGATTTAAGCAACCAAGCTGTGTCTGCGGAAATGACTTC 243
 QY 271 CGAAGACCACTCTTAACCTTAATTTGTGCTCTCAAGACTGAGTGTGCTTGAAT 330
 DB 244 CGAAATGCCACATCTTAACCTTAATTTATGCTTCAAGACTTGAAGTGTGCTTCCTAT 303
 QY 331 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAACTGCTTACCTG 390
 DB 304 CAACACGTTCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGAACTGCTTACCTT 363

[illegible]

Db		1441	CTGAGCGAAGATGATGAGTACTGGGCACCTTTGTCCGACC GGGA CCTTAAGCGG	1500
Oy		1528	AACGACCTGTCTCTGTGGCCAGCTTATAATCTGACCTGACAGTACCTCAAGCTTGACTGG	1587
Db		1501	GAAAGTGTGCTCTGTGGCCAGCCTTACACCAGACCGAGTCAGTACTGAACTGGAATTGG	1560
Oy		1588	AACATGAGCCTCGAANAGATCTCAAAGAACCGCGGTGTGATTTTTTGACCAAGCACCATC	1647
Db		1561	AGGTGAGGTGGGACAGAATCTGAAGAGCAAGAGTGTGATTTTGATGATTAACATT	1620
Oy		1648	CCCC 1651	
Db		1621	GTCC 1624	
 RESULT 15 US-10-451-168-42 Sequence 42, Application US/10451168 Publication No. US20040091969A1 GENERAL INFORMATION: APPLICANT: SMITHKLINE BEECHAM CORPORATION APPLICANT: GLAXO GROUP LIMITED TITLE OF INVENTION: NOVEL COMPOUNDS FILE REFERENCE: GP50039 CURRENT APPLICATION NUMBER: US/10/451,168 PRIOR FILING DATE: 2003-11-12 PRIOR APPLICATION NUMBER: PCT/US01/49232 PRIOR FILING DATE: 2000-12-17 PRIOR APPLICATION NUMBER: 60/256,710 PRIOR FILING DATE: 2000-12-19 PRIOR APPLICATION NUMBER: 60/257,048 PRIOR FILING DATE: 2000-12-20 PRIOR APPLICATION NUMBER: 60/260,482 PRIOR FILING DATE: 2001-01-09 PRIOR APPLICATION NUMBER: 60/264,922 PRIOR FILING DATE: 2001-01-30 PRIOR APPLICATION NUMBER: 60/266,797 PRIOR FILING DATE: 2001-02-06 PRIOR APPLICATION NUMBER: 60/276,988 PRIOR FILING DATE: 2001-03-19 PRIOR APPLICATION NUMBER: 60/281,535 PRIOR FILING DATE: 2001-04-04 PRIOR APPLICATION NUMBER: 60/289,622 PRIOR FILING DATE: 2002-06-28 NUMBER OF SEQ ID NOS: 110 SOFTWARE: FastSeq for Windows Version 4.0 SEQ ID NO 42 LENGTH: 1071 TYPE: DNA ORGANISM: Homo sapiens US-10-451-168-42				
Query Match	58.5%;	Score 1021.4;	DB 18; Length 1071;	
Beat Local Similarity	99.9%;	Pred. No. 0;		
Matches 1022;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Oy		724	ATACTGTCTCCCATGCCCAAGGCTTATTCACAAGGCATCATGAGAGTGGGGTGGCC	783
Db		49	ATACTGTCTCCCATGCCCAAGGCTTATTCACAAGGCATCATGAGAGTGGGGTGGCC	108
Oy		784	ATATCCCTTAACCTGAGGCCCATATATTATGGAAGAGTAGAGACTTGCAAGTGTGGCA	843
Db		109	ATATCCCTTAACCTGAGGCCCATATATTATGGAAGAGTAGAGACTTGCAAGTGTGGCA	168
Oy		844	CATTTCGTGTGAACAATGCGTCAAGCTTGAGGCCCTGCTGAGAGTCCGAGAGACAAA	903
Db		169	CATTTCGTGTGAACAATGCGTCAAGCTTGAGGCCCTGCTGAGAGTCCGAGAGACAAA	228
Oy		904	CCTCCAAAGAGCTGTGAACCTTCAGCCAGAAAACAAAGTCTTTCACCTCGAGTGTGGAT	963
Db		229	CCTCCAAAGAGAGCTGTGAACCTTCAGCCAGAAAACAAAGTCTTTCACCTCGAGTGTGGAT	288

OY 964 GGTCCTTCTTCTTAATGAGCCTCAATATTGTCTCAGAAAGCATTTAAGCAATT 1023
|||||
DB 289 GGTCCTTCTTCTTAATGAGCCTCAATATTGTCTCAGAAAGCATTTAAGCAATT 348
|||||
OY 1024 CCTTCATCATCGAGTCATAATACCAAGTGTGCTTCTGCTGCTATGAAAGAGCT 1083
|||||
DB 349 CCTTCATCATCGAGTCATAATACCAAGTGTGCTTCTGCTGCTATGAAAGAGCT 408
|||||
OY 1084 CCTGAGATTCCTGAGTGTCTCCAAAGTGTGCTTCTGCTGCTATGATAAACAATCTCTG 1143
|||||
DB 409 CCTGAGATTCCTGAGTGTCTCCAAAGTGTGCTTCTGCTGCTATGATAAACAATCTCTG 468
|||||
OY 1144 CACATCCGCGCTCAGTAAATTGACCTTGTGCTAAATGAATACTTCATGACAAGCACTCC 1203
|||||
DB 469 CACATCCGCGCTCAGTAAATTGACCTTGTGCTAAATGAATACTTCATGACAAGCACTCC 528
|||||
OY 1204 CTGACTGAATTCGAGACAGTCTTCTGACTTGTGAGATGTGTCTTGTGTGCTCCT 1263
|||||
DB 529 CTGACTGAATTCGAGACAGTCTTCTGACTTGTGAGATGTGTCTTGTGTGCTCCT 588
|||||
OY 1264 GCACTGATCAAGCTCGATATCAAGAGATGCTGTGCACTGTCTACTTATGAGTTT 1323
|||||
DB 589 GCACTGATCAAGCTCGATATCAAGAGATGCTGTGCACTGTCTACTTATGAGTTT 648
|||||
OY 1324 CGGCACCGGCGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTGTCAAAGCGAGCAAGCT 1383
|||||
DB 649 CGGCACCGGCGCTCAGTGTCTTGAAGACAGAAAGCGGCTTTGTCAAAGCGAGCAAGCT 708
|||||
OY 1384 GATGAAGTCCGCTTGTGTCCGTTGTGCTCTGAAAGGGGACATTTGTATGTTGAA 1443
|||||
DB 709 GATGAAGTCCGCTTGTGTCCGTTGTGCTCTGAAAGGGGACATTTGTATGTTGAA 768
|||||
OY 1444 GAGGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATAAATACTGGGCTACCTTT 1503
|||||
DB 769 GAGGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATAAATACTGGGCTACCTTT 828
|||||
OY 1504 GCTGGAACCGGGAATCTTAATGGGAAAGACCTGTCTGTGGCCAGCTTATATCTGACT 1563
|||||
DB 829 GCTGGAACCGGGAATCTTAATGGGAAAGACCTGTCTGTGTGGCCAGCTTATATCTGACT 888
|||||
OY 1564 GAGCAGTACCTTCAGCTGACTTGAACATGAGCTTCGACAGACTCAAAAGAACCGCGG 1623
|||||
DB 889 GAGCAGTACCTTCAGCTGACTTGAACATGAGCTTCGACAGACTCAAAAGAACCGCGG 948
|||||
OY 1624 GTGATTTTGGACAGACCAATCCCTGATCTGTCTGCTCCGACATGCTCCAGAT 1683
|||||
DB 949 GTGATTTTGGACAGACCAATCCCTGATCTGTCTGCTCCGACATGCTCCAGAT 1008
|||||
OY 1684 CCTCTTTCTTCTTAATCTTCTCTCTCTCCCTCCAGCCCTTTCTTTTCTTTGTGCTCCT 1743
|||||
DB 1009 CCTCTTTCTTCTTAATCTTCTCTCTCTCTCTCCAGCCCTTTCTTTTCTTTGTGCTCCT 1068
|||||
OY 1744 TGA 1746
|||
DB 1069 TGA 1071

Search completed: June 15, 2005, 08:31:11
Job time : 1129.45 secs

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description
1	1746	100.0	1746	6	AR438765	Sequence
2	1746	100.0	1746	6	AX600126	Sequence
3	1746	100.0	2158	6	AR438764	Sequence
4	1746	100.0	2158	6	AX600124	Sequence
5	1656	94.8	2047	9	BC069501	Homo sapi
6	1651.6	94.6	2232	6	AX477711	Sequence
7	1631.2	93.4	2092	6	AX746724	Sequence
8	1631.2	93.4	2092	9	AK090997	Homo sapi
9	1509.2	86.4	2298	9	BC039073	Homo sapi
10	1425	81.6	1716	6	CO733044	Sequence
11	1386.2	79.4	1857	6	AX958392	Sequence
12	1344.4	77.0	1896	6	BC069548	Homo sapi
13	1344.4	77.0	1662	6	AX714052	Sequence
14	1344.4	77.0	1862	9	AK056109	Homo sapi
15	1198.4	68.6	1728	4	AB186392	Canis fam
16	1102	63.1	2139	4	AB045377	Felis cat
17	1041.2	59.6	1728	10	AB186393	Mus muscu
18	951.6	54.5	1739	9	AB060873	Macaca fa
19	634.6	36.3	921	6	AX380871	Sequence

20	413.8	23.7	2228	10	D50580	D50580 <i>Rattus norvegicus</i>
21	408.2	23.4	3391	9	AK095522	AK095522 <i>Homo sapiens</i>
22	406.6	23.3	1846	9	D50579	D50579 <i>Homo sapiens</i>
23	406.6	23.3	2117	6	CQ723952	CQ723952 <i>Sequence</i>
24	406.6	23.3	2191	6	AX330319	AX330319 <i>Sequence</i>
25	406.6	23.3	2191	6	AX330320	AX330320 <i>Sequence</i>
26	406.6	23.3	2191	6	AX331596	AX331596 <i>Sequence</i>
27	406.6	23.3	2191	6	AX411210	AX411210 <i>Sequence</i>
28	406.6	23.3	2191	9	HSIC	Y09616 <i>H. sapiens m</i>
29	406.6	23.3	3155	9	BC032095	BC032095 <i>Homo sapi</i>
30	405	23.2	1748	9	U60553	U60553 <i>Homo sapien</i>
31	401.8	23.0	1831	10	BC028907	BC028907 <i>Mus muscu</i>
32	401.8	23.0	2139	9	BC053670	BC053670 <i>Homo sapi</i>
33	400.2	22.9	2114	9	AK000105	AK000105 <i>Homo sapi</i>
34	400.2	22.9	3824	6	AR528722	AR528722 <i>Sequence</i>
35	400.2	22.9	3824	6	AX358956	AX358956 <i>Sequence</i>
36	400.2	22.9	3824	6	AX362449	AX362449 <i>Sequence</i>
37	400.2	22.9	3824	6	AX454680	AX454680 <i>Sequence</i>
38	400.2	22.9	3824	6	AX464408	AX464408 <i>Sequence</i>
39	400.2	22.9	3824	6	AX491158	AX491158 <i>Sequence</i>
40	400.2	22.9	3824	9	AY358609	AY358609 <i>Homo sapi</i>
41	398.2	22.8	1641	6	AX359653	AX359653 <i>Sequence</i>
42	398.2	22.8	2087	6	AX359663	AX359663 <i>Sequence</i>
43	397.8	22.8	1848	10	D50577	D50577 <i>Meocricetui</i>
44	397.8	22.8	2846	10	BC027185	BC027185 <i>Mus muscu</i>
45	396.2	22.7	2012	10	BC055062	BC055062 <i>Mus muscu</i>

ALIGNMENTS

RESULT 1				
AR438765				
LOCUS	AR438765	1746 bp	DNA	linear
DEFINITION	Sequence 3 from patent US 6664091.			
ACCESSION	AR438765			
VERSION	AR438765.1	GI:42663740		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1746)			
TITLE	Curtis, R.A.V. and Silos-Santiago, I.			
JOURNAL	5310, a human carboxylesterase family member and uses thereof			
FEATURES	Patent: US 6664091-A 3 16-DEC-2003;			
SOURCE	Location/Qualifiers			
	1..1746			

ORIGIN

Query Match	100.0%	Score 1746;	DB 6;	Length 1746;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1746;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCCACAGGGACTTACCTTCATCTGCTTCACAAATGATGCTTTTCTCGATTTCTCCAGCCC	60	
Db	1	ATGCCACAGGGACTTACCTTCATCTGCTTCACAAATGATGCTTTTCTCGATTTCTCCAGCCC	60	
Qy	61	CTGTTGGGACACACACAGTGGGAAAACTGGGCTTCTGGTGAAGGGCCACAGGGAAC	120	
Db	61	CTGTTGGGACACACACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGGGAAC	120	
Qy	121	ACCAAGGCTGGATGATTCAGGGGCAAGCAATCACTGTGCTGGAGGCTGTGCTGTG	180	
Db	121	ACCAAGGCTGGATGATTCAGGGGCAAGCAATCACTGTGCTGGAGGCTGTGCTGTG	180	
Qy	181	AACGTGTTCTCGGAGTCCCTTTGCTGCTCCCGCTGGATCTTCGCAATTACGAAC	240	
Db	181	AACGTGTTCTCGGAGTCCCTTTGCTGCTCCCGCTGGATCTTCGCAATTACGAAC	240	
Qy	241	CCGAGCCTGATATGCCCTTGCGGATTAATTGGAGAAAGCACTCTCAACCAATTGGGC	300	

Db	241	CCGACGCTCGATCCGCCCTGGAGTAACCTTGAGAAAGCACTCTCTACCTTAATTGTGC	300
QY	301	CTTCAGAACTCAGATGGCTGCTCTTAGATCAACATGCTCAAGATGATTAACCGAA	360
Db	301	CTTCAGAACTCAGATGGCTGCTCTTAGATCAACATGCTCAAGATGATTAACCGAA	360
QY	361	TTCCGAGGTCAAGAACTGCTCTTAACCTGAACATCTATGCGCTGCCACGCGGATACA	420
Db	361	TTCCGAGGTCAAGAACTGCTCTTAACCTGAACATCTATGCGCTGCCACGCGGATACA	420
QY	421	GGCTCCAAAGCTCCCGCTTGGTGTGTTCCAGAGAGTGCTTCAAGATCGGTACGC	480
Db	421	GGCTCCAAAGCTCCCGCTTGGTGTGTTCCAGAGAGTGCTTCAAGATCGGTACGC	480
QY	481	TTCACTTTGATGGTCCGCCCTGGCTGCTATGAGAGACGTGCTGTGTGTGTCACAG	540
Db	481	TTCACTTTGATGGTCCGCCCTGGCTGCTATGAGAGACGTGCTGTGTGTGTCACAG	540
QY	541	TACCGGCTAGGAATATTTGGTTTCTTACCAACATGGGATACGATGCTCCGGGAACTGG	600
Db	541	TACCGGCTAGGAATATTTGGTTTCTTACCAACATGGGATACGATGCTCCGGGAACTGG	600
QY	601	GCCTTCAAGAACCAAGTGGCTGCTCTGTCTGGGTCCAGAAAGAACATCGAGTTCTTCGGT	660
Db	601	GCCTTCAAGAACCAAGTGGCTGCTCTGTCTGGGTCCAGAAAGAACATCGAGTTCTTCGGT	660
QY	661	GGGAGCCCAAGCTCTGTGACCATTTTGGGGAGTCCGCGGAGCCATTAATGTTTCTAGT	720
Db	661	GGGAGCCCAAGCTCTGTGACCATTTTGGGGAGTCCGCGGAGCCATTAATGTTTCTAGT	720
QY	721	CTTATACTGTCTCCCATGGCCAAAGCTTATTCACAAAGCCATCATGAGAGTGGGGTG	780
Db	721	CTTATACTGTCTCCCATGGCCAAAGCTTATTCACAAAGCCATCATGAGAGTGGGGTG	780
QY	781	GGCATATCCCTTAACCTGGAGAGCCCATGATTAAGAAAGTGAAGACCTGACGTGGTT	840
Db	781	GGCATATCCCTTAACCTGGAGAGCCCATGATTAAGAAAGTGAAGACCTGACGTGGTT	840
QY	841	GCAATTTCTGTGGTAACATGGGTGAGACTTGAAGGCTGCTGAGAGTGGCTGAGAGCA	900
Db	841	GCAATTTCTGTGGTAACATGGGTGAGACTTGAAGGCTGCTGAGAGTGGCTGAGAGCA	900
QY	901	AAACCCCTCCAAAGAGCTGTGACCCCTCAGCCAGCAAAAACAAAGCTTTTCACTGAGTGGTT	960
Db	901	AAACCCCTCCAAAGAGCTGTGACCCCTCAGCCAGCAAAAACAAAGCTTTTCACTGAGTGGTT	960
QY	961	GATGGTCTTTCTTCTCTAATGAGCTCTAGATCTATTTGTCTCAAGAAACATTTAACA	1020
Db	961	GATGGTCTTTCTTCTCTAATGAGCTCTAGATCTATTTGTCTCAAGAAACATTTAACA	1020
QY	1021	ATTCTTTCCATCATCGGAGTCAATTAACAAGATGGCTTCTGCTGCTATGAGAG	1080
Db	1021	ATTCTTTCCATCATCGGAGTCAATTAACAAGATGGCTTCTGCTGCTATGAGAG	1080
QY	1081	GCTCTGAGATCTCTGAGTGGCTCAACAAAGTCCCTTGGCCCTGCATCTAATCAAAACATC	1140
Db	1081	GCTCTGAGATCTCTGAGTGGCTCAACAAAGTCCCTTGGCCCTGCATCTAATCAAAACATC	1140
QY	1141	CTGCAATCCGCGCTCAAGTATTGGAACCTTGTGGTAAATGATATTTCCATGACAAAGCA	1200
Db	1141	CTGCAATCCGCGCTCAAGTATTGGAACCTTGTGGTAAATGATATTTCCATGACAAAGCA	1200
QY	1201	TCCCTGACCTGAATCCGAGACAGTCTTCTGACCTTGGAGATGTGTTCTTTGTGTGTC	1260
Db	1201	TCCCTGACCTGAATCCGAGACAGTCTTCTGACCTTGGAGATGTGTTCTTTGTGTGTC	1260
QY	1261	CCTCGACTGATCAGCTCGATATCAAGAGTGTGTGACCTGTCTACTTCTTAGAG	1320
Db	1261	CCTCGACTGATCAGCTCGATATCAAGAGTGTGTGACCTGTCTACTTCTTAGAG	1320
QY	1321	TTTGGGACCGGGCCCTCAGTCTTTGAAACAACGAAGCCGCTTTTGTCAAGCCGACAC	1380
Db	1321	TTTGGGACCGGGCCCTCAGTCTTTGAAACAACGAAGCCGCTTTTGTCAAGCCGACAC	1380

[illegible]

Qy	241	CCGACGECTGCATTCGCCCTTGGGATACTTGCAGAAAGCACCTCCTACCTTAATTGTGC	307
Db	336	CCGACACCCTGCATATCGCCCTGGGATTAATTGCGAAGGCACTCTTACCTTAATTGTGC	399
Qy	301	CTCCAGAACTCAGAGAGGCTGCTCTTAGATCAACAAGCTCAAGTGCATTACCAGAA	367
Db	396	CTCCAGAACTCAGAGAGGCTGCTCTTAGATCAACAAGCTCAAGTGCATTACCAGAA	455
Qy	361	TTGCGAGTGTCAAGAAAGCTGCTCTACCTGAAACATATAGCGCTGCGCACGCCGATACA	420
Db	456	TTGCGAGTGTCAAGAAAGCTGCTCTACCTGAAACATATAGCGCTGCGCACGCCGATACA	515
Qy	421	GGCTCAAGGCTCCCGCTCTGT	480
Db	516	GGCTCAAGGCTCCCGCTCTGT	575
Qy	481	TCCATCTTTGATGGGTCCGCGCTGAGTGCCTATAGAGACGTCGTGTGTGTGTGTGTGTGT	540
Db	576	TCCATCTTTGATGGGTCCGCGCTGAGTGCCTATAGAGACGTCGTGTGTGTGTGTGTGTGT	635
Qy	541	TACCGGCTAAGGAATTTTGT	600
Db	636	TACCGGCTAAGGAATTTTGT	695
Qy	601	GCCTTCAAGGACACAGGT	660
Db	696	GCCTTCAAGGACACAGGT	755
Qy	661	GGGAGACCCAGCTCTGT	720
Db	756	GGGAGACCCAGCTCTGT	815
Qy	721	CTTATACGTCTCCCATGAGCCAAAGGCTTATTCACAAAGCCATCATAGAGAGTGGGGTG	780
Db	816	CTTATACGTCTCCCATGAGCCAAAGGCTTATTCACAAAGCCATCATAGAGAGTGGGGTG	875
Qy	781	GCCATCATCCCTTAAGTGAAGGCCCATGATTATAGAAAGTGAAGGACCTGCAAGTGTT	840
Db	876	GCCATCATCCCTTAAGTGAAGGCCCATGATTATAGAAAGTGAAGGACCTGCAAGTGTT	935
Qy	841	GCACATTTCTGTGGTAAACAATGGGTGAGACTGAGAGCCCTGTAGAGTGGCTGAGAGACA	900
Db	936	GCACATTTCTGTGGTAAACAATGGGTGAGACTGAGAGCCCTGTAGAGTGGCTGAGAGACA	995
Qy	901	AAACCTCCAGAGAGCTGTGACCTCAGCCAGCAAAACAAAGTCTTCACTGAGTGGTT	960
Db	996	AAACCTCCAGAGAGCTGTGACCTCAGCCAGCAAAACAAAGTCTTCACTGAGTGGTT	1055
Qy	961	GATGGTGTCTTCTTCTCTAATGAGCCTTATAGATCTATGTCTCAAGAAAGCATTTAAACA	1020
Db	1056	GATGGTGTCTTCTTCTCTAATGAGCCTTATAGATCTATGTCTCAAGAAAGCATTTAAACA	1115
Qy	1021	ATTCTCTTCATCATCCGAGTCAATAACAAGAGTGGCTTCTGTGCTATAGAAAGAG	1080
Db	1116	ATTCTCTTCATCATCCGAGTCAATAACAAGAGTGGCTTCTGTGCTATAGAAAGAG	1175
Qy	1081	GGTCCCGAATCTCTAGTGGCTTCCAAACAAGTCCCTTGCCCTTCATCTGATACAAACATC	1140
Db	1176	GGTCCCGAATCTCTAGTGGCTTCCAAACAAGTCCCTTGCCCTTCATCTGATACAAACATC	1235
Qy	1141	CTGCATATCCCGCTCAGATTAATTTGACCTGTGGCTATAGATCTTCACAGACAGAC	1200
Db	1236	CTGCATATCCCGCTCAGATTAATTTGACCTGTGGCTATAGATCTTCACAGACAGAC	1295
Qy	1201	TCCCTGATCTGAATTCAGACAGTCTTTTGAATTGACTTTGGAGATGTGTCTTTGTGTCT	1260
Db	1296	TCCCTGATCTGAATTCAGACAGTCTTTTGAATTGACTTTGGAGATGTGTCTTTGTGTCT	1355
Qy	1261	CCTGCATGATCAGAGCTCGATATCAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1320
Db	1356	CCTGCATGATCAGAGCTCGATATCAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1415
Qy	1321	TTTGGGACCGGCTCAGTGTCTTTGAAGACAGAAAGCGGCTTTTGTCAAGCGGACAC	1380

Db	1416	TTTCGGACCGGCTCAGTGCCTTGGAAACAGAAAGCCGGCTTTTGTCAAAAGCCGAC	1475
Qy	1381	GCTGATGAAAGTCCGCTTGTGTGTGCGATGATGCTCTCTGAAGGGGGAATATGTAATGTC	1440
Db	1476	GCTGATGAAAGTCCGCTTGTGTGTGCGATGATGCTCTCTGAAGGGGGAATATGTAATGTC	1533
Qy	1441	GAAGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATACTGGGCTACC	1500
Db	1536	GAAGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATACTGGGCTACC	1599
Qy	1501	TTTGCTCGAACCGGGAAATCCTAATGGGGAACGACTGTCTCTGTGTGCCAGCTTAAATCTG	1560
Db	1596	TTTGCTCGAACCGGGAAATCCTAATGGGGAACGACTGTCTCTGTGTGCCAGCTTAAATCTG	1655
Qy	1561	ACTGAGCAGTAACTCTCAGCTGAGCTTGAACATGAGCTCGGACAGAGACTCAAGAAACCG	1620
Db	1656	ACTGAGCAGTAACTCTCAGCTGAGCTTGAACATGAGCTCGGACAGAGACTCAAGAAACCG	1715
Qy	1621	CGGATGATTTTTTGGACACGACCATCCCTGATCTGTGTGTGCTCGACATAGCTCCAC	1680
Db	1716	CGGATGATTTTTTGGACACACACCATCCCTGATCTGTGTGTGCTCGACATAGCTCCAC	1775
Qy	1681	AGTCTCTTCTTCTCTTAACTTTCCTCTCTCTCTCCAGGCTTCTTTTCTTTGTGCT	1740
Db	1776	AGTCTCTTCTTCTTCTTAACTTTCCTCTCTCTCTCCAGGCTTCTTTTCTTTGTGCT	1835
Qy	1741	CTTTGA 1746	
Db	1836	CTTTGA 1841	

Query Match

100.0%; Score 1746; DB 6; Length 2158;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGCCACAGGGAATTACTTATCTGCTTCAAGATGCTTTTCTGATTTCTCCAGCC 60
DB 96 ATGCCACAGGGAATTACTTATCTGCTTCAAGATGCTTTTCTGATTTCTCCAGCC 155
OY 61 CTGTTGGGACACAGCAGTGGGAAAACCTGGGCTTCTGCTGAAGGGCAAGAGAAC 120
DB 156 CTGTTGGGACACAGCAGTGGGAAAACCTGGGCTTCTGCTGAAGGGCAAGAGAAC 215
OY 121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTG 180
DB 216 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGAGCCCTGTGCTGTG 275
OY 181 AAGGTGTTCTGTGGATGCTCCCTTTGTGCTGCCGCTGGAGTCCCTGGATTTACGAAC 240
DB 276 AAGGTGTTCTGTGGATGCTCCCTTTGTGCTGCCGCTGGAGTCCCTGGATTTACGAAC 335
OY 241 CCGGAGCCTGCACTGGCCTGGGATACTTGGAGAAAGCACTCTAACCCTAATTGTGTC 300
DB 336 CCGGAGCCTGCACTGGCCTGGGATACTTGGAGAAAGCACTCTAACCCTAATTGTGTC 395
OY 301 CTCCAGAACTCAGAGTGGCTCTCTTGAATCAACAATGCTCAAGGTGATTAACCGAAA 360
DB 396 CTCCAGAACTCAGAGTGGCTCTCTTGAATCAACAATGCTCAAGGTGATTAACCGAAA 455
OY 361 TTCCGAGTGTCAAGAACTGCTCTTCACTGAACATCTATGCTGCTGCCACCGCAATAC 420
DB 456 TTCCGAGTGTCAAGAACTGCTCTTCACTGAACATCTATGCTGCTGCCACCGCAATAC 515
OY 421 GGCCTCAAGCTCCCGCTCTGTGTGTGTTCCAGAGAGGCTCTTCAAGCTGGCTCAGCC 480
DB 516 GGCCTCAAGCTCCCGCTCTGTGTGTGTTCCAGAGAGGCTCTTCAAGCTGGCTCAGCC 575
OY 481 TCCATCTTGAATGGTCCGCTGCTGCTGCTGAATGAGAGAGTCTGCTGCTGCTCAG 540
DB 576 TCCATCTTGAATGGTCCGCTGCTGCTGCTGAATGAGAGAGTCTGCTGCTGCTCAG 635
OY 541 TACCGGCTAGGAATTTTGTGTTCTTCAACACATGAGATCAGCATGCTCCGGGAACTGG 600
DB 636 TACCGGCTAGGAATTTTGTGTTCTTCAACACATGAGATCAGCATGCTCCGGGAACTGG 695
OY 601 GCCTCAAGAGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 696 GCCTCAAGAGCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
OY 661 GGGGACCCCAAGCTCTGTGACCACTTTTGGCGAGTCCGCGGAGCAATAAGTCTTCTAGT 720
DB 756 GGGGACCCCAAGCTCTGTGACCACTTTTGGCGAGTCCGCGGAGCAATAAGTCTTCTAGT 815
OY 721 CTATATAGTCTCTCCATGAGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTG 780
DB 816 CTATATAGTCTCTCCATGAGCCAAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTG 875
OY 781 GCCATCATCTCTTACCTGAGAGCCCAATGATTTATGAGAGAGTGAAGACTTGCAGTGGT 840
DB 876 GCCATCATCTCTTACCTGAGAGCCCAATGATTTATGAGAGAGTGAAGACTTGCAGTGGT 935
OY 841 GCACATTTCTGTGTAACAATGCTCAAGCTCTGAGGCTGCTGAGGCTCTGAGAGCA 900
DB 936 GCACATTTCTGTGTAACAATGCTCAAGCTCTGAGGCTGCTGAGGCTCTGAGAGCA 995
OY 901 AAACCTCTCAAGAGCTGTGACCTCTCAGCGAGAAACAAAGTCTTCACTGAGTGGT 960
DB 996 AAACCTCTCAAGAGCTGTGACCTCTCAGCGAGAAACAAAGTCTTCACTGAGTGGT 1055
OY 961 GATGTGCTTTTCTTCTTATAGAGCTCTAATCTATTTGTCTCAAGAAAGCATTTTAAAGA 1020
DB 1056 GATGTGCTTTTCTTCTTATAGAGCTCTAATCTATTTGTCTCAAGAAAGCATTTTAAAGA 1115
OY 1021 ATTCCTTCATCATGAGAGTCAATAACAGAGTGTGCTTCTGCTGCTTATGAAGAG 1080

```

```

DB 1116 ATTCCTTCATCATGAGAGTCAATAACAGAGTGTGCTTCTGCTGCTTATGAAGAG 1175
OY 1081 GCTCCTGAGATCCCTAGAGGCTCCCAAGATCCCTTGGCTCCATCATTAACAAACATC 1140
DB 1176 GCTCCTGAGATCCCTAGAGGCTCCCAAGATCCCTTGGCTCCATCATTAACAAACATC 1235
OY 1141 CTGCACATCCCGCTCAGTATTTTGGACCTTGTGCTAATGAATACCTTCATGACAAAGAC 1200
DB 1236 CTGCACATCCCGCTCAGTATTTTGGACCTTGTGCTAATGAATACCTTCATGACAAAGAC 1295
OY 1201 TCCCTGACTGAATTCGAGACAGTCTTCTGACCTTGTGAGATGTGTTCTTTGTGCTC 1260
DB 1296 TCCCTGACTGAATTCGAGACAGTCTTCTGACCTTGTGAGATGTGTTCTTTGTGCTC 1355
OY 1261 CCTGACATGATCAAGCTCGATATTCAGAGATGTGTGTCACCTGTCTAATCTTATGAG 1320
DB 1356 CCTGACATGATCAAGCTCGATATTCAGAGATGTGTGTCACCTGTCTAATCTTATGAG 1415
OY 1321 TTTCCGACACCGGCTCAGTGTCTTTGAAGACAGAGCCGCTTTTGTCAAGCCGACAC 1380
DB 1416 TTTCCGACACCGGCTCAGTGTCTTTGAAGACAGAGCCGCTTTTGTCAAGCCGACAC 1475
OY 1391 GCTGATGAAGTCCGCTTTGTGTTGCTGTGCTCTTCTGAAGGGGACATTTGTTTC 1440
DB 1476 GCTGATGAAGTCCGCTTTGTGTTGCTGTGCTCTTCTGAAGGGGACATTTGTTTC 1535
OY 1441 GAAGAGCCACAGAGAGAGAGAGTACTGAGCCCGGAGAGATGAATATCTGGGCTACC 1500
DB 1536 GAAGAGCCACAGAGAGAGAGAGTACTGAGCCCGGAGAGATGAATATCTGGGCTACC 1595
OY 1501 TTTGCTCGAACCGGGAATCTTATGGAACGACCTGTCTCTGTGAGCCAGCTTATATCTG 1560
DB 1596 TTTGCTCGAACCGGGAATCTTATGGAACGACCTGTCTCTGTGAGCCAGCTTATATCTG 1655
OY 1561 ACTGAGCAGTACCTCAGCTGACCTTGAACATGAGCTCTGGAACAGAGCTCAAGAACCG 1620
DB 1656 ACTGAGCAGTACCTCAGCTGACCTTGAACATGAGCTCTGGAACAGAGCTCAAGAACCG 1715
OY 1621 GGGGTGAGATTTTGGACCAAGCACTCCCTGATCTGTGCTGCTCCGACATGCTTCAC 1680
DB 1716 GGGGTGAGATTTTGGACCAAGCACTCCCTGATCTGTGCTGCTCCGACATGCTTCAC 1775
OY 1681 AGTCTCTTTCTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
DB 1776 AGTCTCTTTCTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1835
OY 1741 CCTTGA 1746
DB 1836 CCTTGA 1841

```

RESULT 5
 BC069501
 LOCUS
 DEFINITION Homo sapiens hypothetical protein FLJ31547, mRNA (cDNA clone
 MGC:97182 IMAGE:7262427), complete cds.
 ACCESSION
 BC069501
 VERSION
 BC069501.1 GI:46854682
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 2047)
 Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Helel,F.,
 Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stappleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Uedin,T.B., Tishitani,S.,
 Carinici,P., Prange,C., Raha,S.S., Logucliano,N.A., Peters,G.J.,
 Abramson,R.D., Mullen,J.S., Bosak,S.A., McEwan,P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.U., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, B., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Buckson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Buetterfeld, Y.S., Krzywinski, M.I., Skalek, U., Smallov, D.E., Scherach, A., Schein, J.B., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2047)
Director MGC Project.
Direct Submission
Submitted (29-APR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Baylor Human Genome Sequencing Center
CDNA Library Preparation: Baylor Human Genome Sequencing Center
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: am@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleaged, H., Kowls, C.R., Snead, A.J., Martin, R.G., Muzny, D.M., Nannavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRBR Place: 4 Row: e Column: 1.
Location/Qualifiers
1..2047
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:97182 IMAGE:7262427"
/tissue_type="PCR rescued clones"
/clone_lib="NIH MGC 277"
/note="Vector: pPCR-Script Amp SK(+) with reversed insert"
1..2047
/gene="PLJ31547"
/db_xref="locusID:221223"
121..1848
/gene="PLJ31547"
/codon_start=1
/product="PLJ31547 protein"
/protein_id="AAH69501.1"
/db_xref="GI:46854683"
/db_xref="locusID:221223"
/translation="MSGNVHPGQILWAIWIAAPTKGSAAGPQNRILGNTGKOVTLAGSPVNVFLGVPFAAPPLGSLRFTNPQASPMVNDLRKATSYFNLCLQSEWTLDDQMLKVAHPKEGVSECLVNTYAPAHADTGSKLVWVFGAATKTSASIPGSAALVAVVVOYRLGIFGFTTDOHAGNMAFDQVAASTWOKNIEFGDPS SVITFGESAGAIIVSRLISPMKGLPHKAIMESGVAITVLAHAYEKSEDDQVAHPCGNADSESLRLCLRTKPSKELLTISQCTSPTRVDAFPFNEPLDLSKAKAIPSTIGVNHCEGELLPMKAPRILISGNSNSLALHILQNTLHPQYLILVNAVYFHDKHSITERDLDLDLGDVFVPPALITAYRHDAQAVYVEYERFRHPOCEDTKPAFVKADHADVRFVFGAFLKGDIVMEGATEEKLSRMMKMYATFARTKPNQNDLSLWPAVNLTEQYLODNLNMSLQRLKEPVDFTSTIPLISASDMLSPSLTFSLTLQPFFFCAP"

ORIGIN

Query Match 94.8%; Score 1656; DB 9; Length 2047;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1656; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 91 GGCCCTTGTCTGTAAGGGCCACAGAGGAACACAGGCTGGATGGAATTCAGGGCAAGCA 150
Db 193 GGCCCTTGTCTGTAAGGGCCACAGAGGAACACAGGCTGGATGGAATTCAGGGCAAGCA 252
Qy 151 GTCACTGTCTGGGAAGCCCTGTGCTGTAACGTGTTCTCGAGAGTCCCTTGTCTGCT 210
Db 253 GTCACTGTCTGGGAAGCCCTGTGCTGTAACGTGTTCTCGAGAGTCCCTTGTCTGCT 312
Qy 211 CCCCCGTGGGATCCCTCGATTTACGAACCCGACGCTGCATGCGCTGGATTAATTG 270
Db 313 CCCCCGTGGGATCCCTCGATTTACGAACCCGACGCTGCATGCGCTGGATTAATTG 372
Qy 271 CGAAGACCACTCCTTACCTAATTGTGCTCCGAACTGAGATGAGTGGCTGCTTGAAT 330
Db 373 CGAAGACCACTCCTTACCTAATTGTGCTCCGAACTGAGATGAGTGGCTGCTTGAAT 432
Qy 331 CAACATGCTCAAGTGCATTTACCCGAAATTCGAGTGCAGAAAGACTGCTTACCTG 390
Db 433 CAACATGCTCAAGTGCATTTACCCGAAATTCGAGTGCAGAAAGACTGCTTACCTG 492
Qy 391 AACATCTATGCGCTGCGCCACGCGATACAGGCTCCAAAGTCCCGCTTGTGTGCTTC 450
Db 493 AACATCTATGCGCTGCGCCACGCGATACAGGCTCCAAAGTCCCGCTTGTGTGCTTC 552
Qy 451 CCAGAGGCTGCTTCAAGACTGCTCAGCTCCATCTTGTATGAGTCCGCTGCTGCGC 510
Db 553 CCAGAGGCTGCTTCAAGACTGCTCAGCTCCATCTTGTATGAGTCCGCTGCTGCGC 612
Qy 511 TATGAGGACGTCGT 570
Db 613 TATGAGGACGTCGT 672
Qy 571 ACATGGATCAGCATGCTCCGAGGAACTGGGCTTCAAGAGCCAGAGTGGCTGCTGTC 630
Db 673 ACATGGATCAGCATGCTCCGAGGAACTGGGCTTCAAGAGCCAGAGTGGCTGCTGTC 732
Qy 631 TGGGTCCAGAAAGCAATCCGATCTTCCGTTGGGAGCCCAAGCTCTGTGACCATCTT 690
Db 733 TGGGTCCAGAAAGCAATCCGATCTTCCGTTGGGAGCCCAAGCTCTGTGACCATCTT 792
Qy 691 GAGTCGCGGAGCCCAATGAAGTGTTCATGCTATATGCTGCTCCATGCGCAAGGCTTA 750
Db 793 GAGTCGCGGAGCCCAATGAAGTGTTCATGCTATATGCTGCTCCATGCGCAAGGCTTA 852
Qy 751 TTCCACAAAGCCATCATGGAAGTGGGTGGCCATCATCCTTACCTGAGGCCCATGAT 810
Db 853 TTCCACAAAGCCATCATGGAAGTGGGTGGCCATCATCCTTACCTGAGGCCCATGAT 912
Qy 811 TATGAGAGTGAAGGACCTGAGAGTGTGTGACATTTCTGTGTGAACAATGCTCAGAC 870
Db 913 TATGAGAGTGAAGGACCTGAGAGTGTGTGACATTTCTGTGTGAACAATGCTCAGAC 972
Qy 871 TCTGAGGCTTCTGTAAGTGTCTGAGAGCAAAACCTCCAAAGAGCTGCTGACCTCAGC 930
Db 973 TCTGAGGCTTCTGTAAGTGTCTGAGAGCAAAACCTCCAAAGAGCTGCTGACCTCAGC 1032
Qy 931 CAGAAACCAAGCTTTCACTGAGTGTGATGAGTGTCTTCTTCTTAATGAGCTCTTA 990
Db 1033 CAGAAACCAAGCTTTCACTGAGTGTGATGAGTGTCTTCTTCTTAATGAGCTCTTA 1092
Qy 991 GATCTATTGCTCAGAAAGCAATTTAAAGCAATCTCTTCATCATCGAGTCAATTAACAC 1050
Db 1093 GATCTATTGCTCAGAAAGCAATTTAAAGCAATCTCTTCATCATCGAGTCAATTAACAC 1152
Qy 1051 GAGTGTGCTTCTGCTGCTCATGAGAGGCTCTGAGATCTCAGTGGCTCCAAAG 1110
Db 1153 GAGTGTGCTTCTGCTGCTCATGAGAGGCTCTGAGATCTCAGTGGCTCCAAAG 1212
Qy 1111 TCCCTTGGCTTCATCTGATACAAAGCAATCTCGACATCCCGCTGATATTGACACTT 1170
Db 1213 TCCCTTGGCTTCATCTGATACAAAGCAATCTCGACATCCCGCTGATATTGACACTT 1272

Qy	1171	GTGCTAATGAATCTTCCATGACGACAGACCTCCGTAATCGGACAGCTTCTG	1230
Db	1273	GTGCTAATGAATCTTCCATGACGACAGCTCCGTAATCGGACAGCTTCTG	1332
Qy	1231	GACTTGCTGGAAGATGTGTTCTTTGTGATCCCTGACATGATCAGCTCGATATCACA	1290
Db	1333	GACTTGCTGGAAGATGTGTTCTTTGTGATCCCTGACATGATCAGCTCGATATCACA	1392
Qy	1291	GATCTGATGACCTGTTCTACTTCTATGATTTTGACACCGAGCTCAAGCTTTGAGAC	1350
Db	1393	GATCTGATGACCTGTTCTACTTCTATGATTTTGACACCGAGCTCAAGCTTTGAGAC	1452
Qy	1351	ACGAAGCCGGCTTTTGTGCAAGCCGACACGCTATGAAATCCGCTTTGTGTGGTGTG	1410
Db	1453	ACGAAGCCGGCTTTTGTGCAAGCCGACACGCTATGAAATCCGCTTTGTGTGGTGTG	1512
Qy	1411	GCCTTCCCTGAAGGGGGACATTTGTTATGTTTCGAAGGACCGAGAGAGAAATGATG	1470
Db	1513	GCCTTCCCTGAAGGGGGACATTTGTTATGTTTCGAAGGACCGAGAGAGAAATGATG	1572
Qy	1471	AGCCGGAAGATGATGAATTAATCTGAGCTTGTGCTGCAACCGGAAATCTTAATGGAA	1530
Db	1573	AGCCGGAAGATGATGAATTAATCTGAGCTTGTGCTGCAACCGGAAATCTTAATGGAA	1632
Qy	1531	GACTGTCTCTGTGAGCCAGCTTATATATGACTAGACAGTACTCTTCACTGATGAA	1590
Db	1633	GACTGTCTCTGTGAGCCAGCTTATATATGACTAGACAGTACTCTTCACTGATGAA	1692
Qy	1591	ATGAGCCTCGGACAGAGACTGAAGAAACCGGGGGGATTTTGGACGAGCAACATCC	1650
Db	1693	ATGAGCCTCGGACAGAGACTGAAGAAACCGGGGGGATTTTGGACGAGCAACATCC	1752
Qy	1651	CTGATCTGTCTGCTCCGACATGCTCCAGAGTCTCTTCTTCTTAACTTCTCTCT	1710
Db	1753	CTGATCTGTCTGCTCCGACATGCTCCAGAGTCTCTCTTCTTCTTAACTTCTCTCT	1812
Qy	1711	CTCCTCCAGCTTCTTCTTCTTGTGTGCTCTGA	1746
Db	1813	CTCCTCCAGCTTCTTCTTCTTGTGTGCTCTGA	1848
RESULT 6			
AX477711		2232 bp	DNA linear PAT 12-AUG-2002D
LOCUS	AX477711		
DEFINITION	Sequence 23 from Patent WO246426.		
ACCESSION	AX477711		
VERSION	AX477711.1 GI:22216866		
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Sanjanwala,M.M., Yao,M.G., Au-Young,J., Baughn,M.R., Arvizu,C., Ring,H.Z., Lee,B.A., Ding,L., Hafalla,A.J., Tang,Y.T., Yue,H., Tribouley,C.M., Lu,D.A., Lal,P.G., Warren,B.A., Yang,J., Walla,N.K., Nguyen,D.B., Gandhi,A.R. and Ison,C.H.		
JOURNAL	Drug metabolizing enzymes		
FEATURES	Patent: WO 0246426-A 23 13-JUN-2002;		
source	Incyte Genomics, Inc. (US)		
	Location/Qualifiers		
	1..2232		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 6538080CB1"		
ORIGIN			
Query Match	94.6%;	Score 1651.6;	DB 6; Length 2232;
Best Local Similarity	99.5%;	Pred. No. 0;	
Matches 1657; Conservative	0;	Mismatches 9;	Indels 0; Gaps 0;
81 GGGAAAACTGGGCTTCTGCTGAAGGCCACACAGAAACACAGCTGGAGTGAATTCA	140		

Db	264	GGTCCAAAGAGAGGGCTTCTGTGAAGGGCCACAGAGGAACACAGGCTGGGATGGATTCA	323
Qy	141	GGGCAAGCAAGTCACTGTGCTGGAGACCTGTGCTGTGAACSTGTTCTCGAGTCCC	200
Db	324	GGGCAACCAAGTCACTGTGCTGGAGAACCCGTGTGCTGTGAACSTGTTCTCGAGTCCC	383
Qy	201	CTTTGCTGCTCCCCCGCTGGGATGCCCTGGCATTTTACGAACCCGAGGCTGATGSCCTG	260
Db	384	CTTTGCTGCTCCCCCGCTGGGATGCCCTGGCATTTTACGAACCCGAGGCTGATGSCCTG	443
Qy	261	GGATTAACCTTGGAGAAACCACTCTACCCCAATTTTGTGCTCAGAACTCAGAGTGGCT	320
Db	444	GGATTAACCTTGGAGAAACCACTCTACCCCAATTTTGTGCTCAGAACTCAGAGTGGCT	503
Qy	321	GCTTGTAGATCAACACATGCTCAAGGTGCATTACCCGAAATTGGAGTGTCAAGAACTG	380
Db	504	GCTTGTAGATCAACACATGCTCAAGGTGCATTACCCGAAATTGGAGTGTCAAGAACTG	563
Qy	381	CCTTACCTTGAACATCTATGCGCTGTGCCCAAGCCGATACAGGCTTCAAAGCTCCCCGT	440
Db	564	CCTTACCTTGAACATCTATGCGCTGTGCCCAAGCCGATACAGGCTTCAAAGCTCCCCGT	623
Qy	441	GGTGTGGTCCACAGAGGTGCTTCAAGCTGAGCTCAATCTTTGATGGGTCCG	500
Db	624	GGTGTGGTCCACAGAGGTGCTTCAAGCTGAGCTCAATCTTTGATGGGTCCG	683
Qy	501	CTTGCGTGCCTATGAGACGTGCTGATTGTGTGTGCAGTACCGGCTAGGAATATTGG	560
Db	684	CTTGCGTGCCTATGAGACGTGCTGATTGTGTGTGCAGTACCGGCTAGGAATATTGG	743
Qy	561	TTTCTTCAACCATAGGAATCAGATGCTCCGCGGGAACCTGGGCTTCAAGAACAGAGTGGC	620
Db	744	TTTCTTCAACCATAGGAATCAGATGCTCCGCGGGAACCTGGGCTTCAAGAACAGAGTGGC	803
Qy	621	TGCTCTGTCCCGGGTCCAGAAAGCATCGAGTTCTTGCGTGGGGACCCAGCTCTGTGAC	680
Db	804	TGCTCTGTCCCGGGTCCAGAAAGCATCGAGTTCTTGCGTGGGGACCCAGCTCTGTGAC	863
Qy	681	CATCTTTGGCAGGTCCGCGGGAGCCATAGTGTCTTAGTCTTTACTGTCTCCCATGGC	740
Db	864	CATCTTTGGCAGGTCCGCGGGAGCCATAGTGTCTTAGTCTTTACTGTCTCCCATGGC	923
Qy	741	CAAAGGCTTATTCCAAAGCCATCATGAGAGGTGGGCTGGCCATCATCCCTTACCTGGA	800
Db	924	CAAAGGCTTATTCCAAAGCCATCATGAGAGGTGGGCTGGCCATCATCCCTTACCTGGA	983
Qy	801	GGCCCATGATTTATAGAAAGGTGAGACCTGACAGTGGTTGCAATTTCTGTGGTAAACA	860
Db	984	GGCCCATGATTTATAGAAAGGTGAGACCTGACAGTGGTTGCAATTTCTGTGGTAAACA	1043
Qy	861	TGCGTCAAGACTCTGAGGCGCTGCTGAGAGTGGCTGAGACAACAAACCTTCAAGAGAGTCT	920
Db	1044	TGCGTCAAGACTCTGAGGCGCTGCTGAGAGTGGCTGAGACAACAAACCTTCAAGAGAGTCT	1103
Qy	921	GACCCCTCAGCCAGAAAACAAAGTCTTTCACCTCGAGTGGTGAATGAGTGTCTTCTCTTAA	980
Db	1104	GACCCCTCAGCCAGAAAACAAAGTCTTTCACCTCGAGTGGTGAATGAGTGTCTTCTCTTAA	1163
Qy	981	TGAGCCTCTAGATTATTTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATCGAGT	1040
Db	1164	TGAGCCTCTAGATTATTTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATCGAGT	1223
Qy	1041	CAATTAACACAGAGTGGCTTCTGCTGCTCATGAGAGAGGCTCTGAGATCTCCAGTGG	1100
Db	1224	CAATTAACACAGAGTGGCTTCTGCTGCTCATGAGAGAGGCTCTGAGATCTCCAGTGG	1283
Qy	1101	CTTCAACAAGTCCCTTGCCCTTCATCTGATACAAAACATCTCTGACATCCCCTTCAGTA	1160
Db	1284	CTTCAACAAGTCCCTTGCCCTTCATCTGATACAAAACATCTCTGACATCCCCTTCAGTA	1343
Qy	1161	TTTTCACCTTGTGGCTAATGAATCTTCATGACAAACATCTCCCTGACTGAAATCCGAGA	1220

Db 1344 TTGCACTTGTGGCTAATGAATTAATTCATGACAAGCACTCCGACTGAATCCGGA 1403
Qy 1221 CAGTCTTGTGAATTGCTTGGAGATGTTCTTTTGTGCTCCGACATGATCAGCTCG 1280
Db 1404 CAGTCTTGTGAATTGCTTGGAGATGTTCTTTTGTGCTCCGACATGATCAGCTCG 1463
Qy 1281 ATATCAGAGATGCTGTGTCACCTGTCTACTTATAGATTGGGACCGGCTCAGTG 1340
Db 1464 ATATCAGAGATGCTGTGTCACCTGTCTACTTATAGATTGGGACCGGCTCAGTG 1523
Qy 1341 CTTTGAAGCAGAGACCGGCTTTTGTCAAGCCGACCACTGATGAAATGCTGTTGT 1400
Db 1524 CTTTGAAGCAGAGACCGGCTTTTGTCAAGCCGACCACTGATGAAATGCTGTTGT 1583
Qy 1401 GTTGGTGTGCTCTCTGAAAGGGGACATTTATTTATGTAAGAGACCGAGAGGA 1460
Db 1584 GTTGGTGTGCTCTCTGAAAGGGGACATTTATTTATGTAAGAGACCGAGAGGA 1643
Qy 1461 GAATTTACTGAGCCGGAAAGATGATAAATCTGGGCTTACCTTTGCTGAAACGGGAATCC 1520
Db 1644 GAATTTACTGAGCCGGAAAGATGATAAATCTGGGCTTACCTTTGCTGAAACGGGAATCC 1703
Qy 1521 TAATGGAAACGACCTGTCTGTGAGCAGCTTAATCTGAAGAGATCTCCAGCT 1580
Db 1704 TAATGGAAACGACCTGTCTGTGAGCAGCTTAATCTGAAGAGATCTCCAGCT 1763
Qy 1581 GGAATTGAACATGAGCTCGAGACAGACTCAAGAACCGGGGTGGAATTTTGAACAG 1640
Db 1764 GGAATTGAACATGAGCTCGAGACAGACTCAAGAACCGGGGTGGAATTTTGAACAG 1823
Qy 1641 CACCATCCGCTGATCTGTGCTCCGACATGCTCCACAGTCTTTTCTTAAAC 1700
Db 1824 CACCATCCGCTGATCTGTGCTCCGACATGCTCCACAGTCTTTTCTTAAAC 1883
Qy 1701 TTTCTCTCTCTCTCTCCAGCTTTCTTTTCTTTTGTGCTCTGA 1746
Db 1884 TTTCTCTCTCTCTCTCCAGCTTTCTTTTCTTTTGTGCTCTGA 1929

RESULT 7
AX746724 2092 bp mRNA linear PAT 20-JUN-2003
LOCUS AX746724
DEFINITION Sequence 249 from Patent EPI308459.
ACCESSION AX746724
VERSION AX746724.1 GI:32130991
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1
Isoqai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tanechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
Maunho, Y.
TITLE
Full-length cDNA sequences
Patent: EP 1308459-A 249 07-MAY-2003;
JOURNAL
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)
FEATURES
Source
1. 2092
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 93.4%; Score 1631.2; DB 6; Length 2092;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 3; Indels 10; Gaps 1;
Qy 91 GGGCTTCTCTGAGAGGACACAGAGAACACCAAGGCTGGATGATTCAGGGCAAGCA 150
Db 138 GGGCTTCTCTGAGAGGACACAGAGAACACCAAGGCTGGATGATTCAGGGCAAGCA 197

Qy 151 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTTTCTCTGAGATCCCTTGTCT 210
Db 198 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTTTCTCTGAGATCCCTTGTCTCT 257
Qy 211 CCCCCCTGGGATCCCTGTGATTTACGAACCCGACGCTGATCGCCCTGGGATTAATCTG 270
Db 258 CCCCCCTGGGATCCCTGTGATTTACGAACCCGACGCTGATCGCCCTGGGATTAATCTG 317
Qy 271 CGAAGAGCAGCTCTTACCTTAATTT-----GTGCTCCAGAACTCAGAGTGCT 320
Db 318 CGAAGAGCAGCTCTTACCTTAATTTGTAGAACAGTGCTCCAGAACTCAGAGTGCT 377
Qy 321 GCTTTAGATCAACATGCTCAAGTGCTATTAACCCGAATTCGAGTGTGTAAGACTG 380
Db 378 GCTTTAGATCAACATGCTCAAGTGCTATTAACCCGAATTCGAGTGTGTAAGACTG 437
Qy 381 CCTTTACCTGAACATCTATGCGCTGCGCAAGCGGATACAGGCTCCAGCTCCCGCTT 440
Db 438 CCTTTACCTGAACATCTATGCGCTGCGCAAGCGGATACAGGCTCCAGCTCCCGCTT 497
Qy 441 GGTGTGTTCAGAGAGGTGCTTCAAGATGAGTCAAGCTCCATCTTTATGAGTCCGC 500
Db 498 GGTGTGTTCAGAGAGGTGCTTCAAGATGAGTCAAGCTCCATCTTTATGAGTCCGC 557
Qy 501 CTTGGCTGCTATGAGGACGTGCTGTTGTGTCGTCAATACCGGCTAGAAATATTGG 560
Db 558 CTTGGCTGCTATGAGGACGTGCTGTTGTGTCGTCAATACCGGCTAGAAATATTGG 617
Qy 561 TTTCTTACCAACATGGGATGAGCATGCTCCGGGAACTGGGCTTCAAGACAGGTGC 620
Db 618 TTTCTTACCAACATGGGATGAGCATGCTCCGGGAACTGGGCTTCAAGACAGGTGC 677
Qy 621 TGTCTGTCTGTGAGTCCAGAAACATCGAGTCTTGGGAGGAGCCGAGCTGTGAC 680
Db 678 TGTCTGTCTGTGAGTCCAGAAACATCGAGTCTTGGGAGGAGCCGAGCTGTGAC 737
Qy 681 CATCTTGGGAGTCCCGGGAGCCATTAAGTGTCTTAATCTGTCTCCATGAC 740
Db 738 CATCTTGGGAGTCCCGGGAGCCATTAAGTGTCTTAATCTGTCTCCATGAC 797
Qy 741 CAAAGCTTATTCACAAAGACCATGAGAGTGGGAGGCAATCCCTTACTGGA 800
Db 798 CAAAGCTTATTCACAAAGACCATGAGAGTGGGAGGCAATCCCTTACTGGA 857
Qy 801 GGCCCATGATTTATGAGAAAGATGAGACCTGAGGTGTGCAATTTCTGTGTAACA 860
Db 858 GGCCCATGATTTATGAGAAAGATGAGACCTGAGGTGTGCAATTTCTGTGTAACA 917
Qy 861 TGGGTGAGCTCTGAGGCTCTGAGGTGCTGAGCAAAACCTTCAGAGAGTCT 920
Db 918 TGGGTGAGCTCTGAGGCTCTGAGGTGCTGAGCAAAACCTTCAGAGAGTCT 977
Qy 921 GACCTCAGCCGAAACAAAGCTTTTCACTGAGTGTGATGATGATCTTTCTTAA 980
Db 978 GACCTCAGCCGAAACAAAGCTTTTCACTGAGTGTGATGATGATCTTTCTTAA 1037
Qy 981 TGAGCTCTAGATCATTTGTCTCAGAAAGCATTTTAAGCAATTCCTTCATCATGAGT 1040
Db 1038 TGAGCTCTAGATCATTTGTCTCAGAAAGCATTTTAAGCAATTCCTTCATCATGAGT 1097
Qy 1041 CAATACAGAGAGTGTGCTTCTGTGCTGATGAGAGGCTCTGAGATCCGAGTGG 1100
Db 1098 CAATACAGAGAGTGTGCTTCTGTGCTGATGAGAGGCTCTGAGATCCGAGTGG 1157
Qy 1101 CTTCAAGAGTCCCTTGCCTTCATCTGATACAAACATCTTCACATCCCGCTCAGTA 1160
Db 1158 CTTCAAGAGTCCCTTGCCTTCATCTGATACAAACATCTTCACATCCCGCTCAGTA 1217
Qy 1161 TTTGACCTTGTGCTATGAAATCTTCCATGACAAAGCACTCCCTGATGAAATCCGAGA 1220
Db 1218 TTTGACCTTGTGCTATGAAATCTTCCATGACAAAGCACTCCCTGATGAAATCCGAGA 1277

QY 1221 CAGTCTTGGACCTTGGAGATGTGTCTTGTGGTCCCTGCACTGATCAGACTCG 1280
 DB 1278 CAGTCTTGGACCTTGGAGATGTGTCTTGTGGTCCCTGCACTGATCAGACTCG 1337
 QY 1281 ATATCAGAGAGTGTGTGTCACCTGTCTACTTATGATGTTGGGACCCGGCTCACTG 1340
 DB 1338 ATATCAGAGAGTGTGTGTCACCTGTCTACTTATGATGTTGGGACCCGGCTCACTG 1397
 QY 1341 CTTTGAACACGAGCCGGCTTTTGTCAAGCCGACGCTGATGTAAGTCCGCTTGT 1400
 DB 1398 CTTTGAACACGAGCCGGCTTTTGTCAAGCCGACGCTGATGTAAGTCCGCTTGT 1457
 QY 1401 GTTCGGTGTGCTCTCTGAAGGGGACATTTTATGTTTCAAGAGCCACGAGAGAGA 1460
 DB 1458 GTTCGGTGTGCTCTCTGAAGGGGACATTTTATGTTTCAAGAGCCACGAGAGAGA 1517
 QY 1461 GAACTTACTGAGCCGGAAGATGATGTAATACTGGGCTACCTTTGCTGAAACCGGGAATCC 1520
 DB 1518 GAACTTACTGAGCCGGAAGATGATGTAATACTGGGCTACCTTTGCTGAAACCGGGAATCC 1577
 QY 1521 TAATGGGAACGACCTGTCTGTGGCCGCTATATATGATGACGAGTACCTCAGCT 1580
 DB 1578 TAATGGGAACGACCTGTCTGTGGCCGCTATATATGATGACGAGTACCTCAGCT 1637
 QY 1581 GGACTTGAACATGAGCTCTGACAGACGCTCAAGAACCGCGGGTGGATTTTGGACCG 1640
 DB 1638 GGACTTGAACATGAGCTCTGACAGACGCTCAAGAACCGCGGGTGGATTTTGGACCG 1697
 QY 1641 CACCAATCCCTGATCTGTCTGCTCCGACATGCTCCACAGTCTCTTTCTTCTTAAC 1700
 DB 1698 CACCAATCCCTGATCTGTCTGCTCCGACATGCTCCACAGTCTCTTTCTTCTTAAC 1757
 QY 1701 TTTCT 1746
 DB 1758 TTTCT 1803

RESULT 8
 LOCUS AK090997
 DEFINITION Homo sapiens cDNA FLJ33678 f18, clone BRAHM2002191, weakly similar to FAT2 ACYL-CoA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).
 ACCESSION AK090997
 VERSION AK090997.1 GI:21749268
 KEYWORDS oiligo capping; f18 (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakematsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, D., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosotani, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Niimiyama, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Horiuchi, T., Kusanagi, T., Kanohori, K., Takahashi, Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, Y., Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kunigami, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiroo, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,

Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hara, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Maehuo, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohtsuka, O., Isoigai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)
 14702039
 2 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakematsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakami, K., Kanohori, K., Takahashi, Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Maehuo, Y., Nagai, K. and Isoigai, T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 2092)
 Isoigai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isoigai, FLJ Project (HRI Team) ; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) ; cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 Location: Qualifiers
 1..2092
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRAHM2002191"
 /tissue_type="brain"
 /clone_id="BRAHM2"
 /note="Cloning vector: PME18SFL3"
 394..1803
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC03565.1"
 /db_xref="GI:21749269"
 /translation="MLKVHPKGVSEDCLYLNTYAPAHADTGSKLPLYVPPGAFK TGSATFPDGSALAEVDYLVVVOYRIGIPGFPTTMDHAPGNVAFKQVVALASVOK NIERFGDPRSSVTI FGSAGAISSYSLISFMARGAFKAIMESGVALIPLYLAEVDE KSEBDQVVAHFCGNASDSEALRLCKRTKSEKILLTSQKTSFTRVVDGAFPRNEPL DLVQAKFKALPSLIGVNSHCGFLPLPKKAPKELISGNSKSLAHLIDNLIHPQYL HLVANEFYHDSGLTEIRDSLDLDIGVFPVPLIATRYHRDAGAPVYFVFRPRPO CPEDTKPAFYADNADYRFFVFGAFPLKGDIVMEGATIESEKRLSRKMKWATFART GPNQGNLSLMPAYNTQYQLDLNLSLGRRLKRPREFMTSTIPLLSMDMLHSP LSLITFSLDLPFFFCAP"

FEATURES

source

CDS

ORIGIN

Query Match 93.4%; Score 1631.2; DB 9; Length 2092;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 3; Indels 10; Gaps 1;
 QY 91 GGGCTTGTGCTGAAGGCCACAGAGAACACGAGCTGGAGTTCAGGGCAAGCA 150
 DB 138 GGGCTTGTGCTGAAGGCCACAGAGAACACGAGCTGGAGTTCAGGGCAAGCA 197
 QY 151 GTCACTGTGCTGGGAAGCCGCTGTGCTGAACGCTTCTCGAGTCCCTTGTCTCT 210
 DB 198 GTCACTGTGCTGGGAAGCCGCTGTGCTGAACGCTTCTCGAGTCCCTTGTCTCT 257
 QY 211 CCCCCGCTGGATCCCTCGATTTACGAACCCGAGCTGATGCGCTCGGAGTAATCTG 270

D	258	CCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCGTGCATGCGCTGGGATTAACCTG	317
Q	271	CGAAGAGCACTCCCTAACCCTAATTT-----GTGCTCCGAATCTCAGAGTGGCT	320
D	318	CGAAGAGCACTCCCTAACCCTAATTTGTAAGACGAGGTGCTCCGAGAACTCAGAGTGGCT	377
Q	321	GCTTTAATCAACATGCTTCAGAGTCAATTAACCGAAATTCGGAAGTCAAGAACTG	380
D	378	GCTTTAATCAACATGCTTCAGAGTCAATTAACCGAAATTCGGAAGTCAAGAACTG	437
Q	381	CCTCAACGTAACATTTATGCGCTGCGCCGAGATTCAGAGTCCAACTCCCGCTT	440
D	438	CCTCAACGTAACATTTATGCGCTGCGCCGAGATTCAGAGTCCAACTCCCGCTT	497
Q	441	GGTGTGTTCAGAGAGTGGCTTCAAGAATGAGTCCAGCTCCATCTTTATAGAGTCCG	500
D	498	GGTGTGTTCAGAGAGTGGCTTCAAGAATGAGTCCAGCTCCATCTTTATAGAGTCCG	557
Q	501	CTTGGCTGCTTAATAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	560
D	558	CTTGGCTGCTTAATAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	617
Q	561	TTTCTTCAACATATGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	620
D	618	TTTCTTCAACATATGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	677
Q	621	TGCTGT	680
D	678	TGCTGT	737
Q	681	CATCTTTGGGAGATCGCGCGGAGCCATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	740
D	738	CATCTTTGGGAGATCGCGCGGAGCCATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	797
Q	741	CAAAAGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCTTAACTGGA	800
D	798	CAAAAGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCTTAACTGGA	857
Q	801	GGCCCATGATTAATGAGAGATGAGAGATCTGAGAGTGTGTGTGTGTGTGTGTGTGTGT	860
D	858	GGCCCATGATTAATGAGAGATGAGAGATCTGAGAGTGTGTGTGTGTGTGTGTGTGTGT	917
Q	861	TGCGTCAAGCTGTGAGGCTGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	920
D	918	TGCGTCAAGCTGTGAGGCTGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	977
Q	921	GACCTTCAAGCCGAGAAACAAAGTCTTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGT	980
D	978	GACCTTCAAGCCGAGAAACAAAGTCTTTCACTGAGTGTGTGTGTGTGTGTGTGTGTGT	1037
Q	981	TGAGCTCTAGATCTATGTTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGT	1040
D	1038	TGAGCTCTAGATCTATGTTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGT	1097
Q	1041	CAATTAACAGAGT	1100
D	1098	CAATTAACAGAGT	1157
Q	1101	CTTCAACAGAGT	1160
D	1158	CTTCAACAGAGT	1217
Q	1161	TTTGGACCTTGT	1220
D	1218	TTTGGACCTTGT	1277
Q	1221	CAGCTTCTGGAATGTTGT	1280
D	1278	CAGCTTCTGGAATGTTGT	1337
Q	1281	ATAATCAGAGATGCTGT	1340
D	1338	ATAATCAGAGATGCTGT	1397

Q	1341	CTTTGAAGACAGAGGCGGCTTTTGTCAAAACCGACCAAGCTGATGTAAGTCCGTTGT	1400
D	1398	CTTTGAAGACAGAGGCGGCTTTTGTCAAAACCGACCAAGCTGATGTAAGTCCGTTGT	1457
Q	1401	GTTCCGT	1460
D	1458	GTTCCGT	1517
Q	1461	GAAAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTACTTTGTCTGAACCGGAAATCC	1520
D	1518	GAAAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTACTTTGTCTGAACCGGAAATCC	1577
Q	1521	TAAATGGAACGACCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1580
D	1578	TAAATGGAACGACCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1637
Q	1581	GGACTTTGAACATGAGCCCTCGACAGAGATCAAAAGAACCGGAGGTGAGTTTGGACAG	1640
D	1638	GGACTTTGAACATGAGCCCTCGACAGAGATCAAAAGAACCGGAGGTGAGTTTGGACAG	1697
Q	1641	CACCATCCCTGATTCCTGT	1700
D	1698	CACCATCCCTGATTCCTGT	1757
Q	1701	TTTCTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGTCTCTTGA	1746
D	1758	TTTCTCTCTCTCTCTCTCAAGCTTTCTTTTCTTTTGTGTCTCTTGA	1803

RESULT 9
 BC039073
 LOCUS
 DEFINITION
 Homo sapiens, similar to carboxylesterase, clone IMAGE:5174937,
 mRNA.
 ACCESSION
 BC039073
 VERSION
 BC039073.1 GI:25058132
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 1 (bases 1 to 2298)
 AUTHORS
 Strausberg, R.
 TITLE
 Direct Submision
 JOURNAL
 Submitted (01-NOV-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 51 Row: K Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 Location/Qualifiers
 1..2298
 /organism="Homo sapiens"

FEATURES
 source

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_image="5174937"
 /issue_type="Brain, Lung, Testis, adult, pooled whole"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN
 Query Match 86.4%; Score 1509.2; DB 9; Length 2298;
 Beel Local Similarity 92.4%; Pred. No. 0;
 Matches 1653; Conservative 0; Mismatches 3; Indels 132; Gaps 1;

91 GGGCTCTGCTGGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGGCAAGCA 150
 103 GGGCTCTGCTGGAAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGGCAAGCA 162
 151 GTCACTGTCTGGGAAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTTGCTGCT 210
 163 GTCACTGTCTGGGAAGCCCTGTGCTGTGAACGTGTTCTTCGAGTCCCTTTGCTGCT 222
 211 CCCCCCTGGGATCCCTGGGATTTACGAACCCGACGCTGCATCCGCTGGGATTAATTG 270
 223 CCCCCCTGGGATCCCTGGGATTTACGAACCCGACGCTGCATCCGCTGGGATTAATTG 282
 271 CGAGAGGCACTCTCACTCACTTAATTTGCTGCTCAAGACTCAGAGTGGCTGCTTAAGT 330
 283 CGAGAGGCACTCTCACTCACTTAATTTGCTGCTCAAGACTCAGAGTGGCTGCTTAAGT 342
 331 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGGTGCAGAGACTGCTCTTAAGT 390
 343 CAACACATGCTCAAGGTGATTAACCGAAATTCGAGGTGCAGAGACTGCTCTTAAGT 402
 391 AACATCTATGCGCTGCTGCCACGCGGATACAGGCTCAAGCTCCCGTCTTGGTGGTTC 450
 403 AACATCTATGCGCTGCTGCCACGCGGATACAGGCTCAAGCTCCCGTCTTGGTGGTTC 462
 451 CCAGAGGCTCTCAAGCTGCTCAAGCTGCTCACTTTGATGGTCCGCTGGCTGCTC 510
 463 CCAGAGGCTCTCAAGCTGCTCAAGCTGCTCACTTTGATGGTCCGCTGGCTGCTC 522
 511 TATGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 569
 523 TATGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 582
 570 ----- 569
 583 TGAATCTCTCAATATGATGATCCAGATCCCAAGAGGCGAGAGTGCAGTACCATTA 642
 570 ----- 569
 643 GACAGTCCCGCAACCTTAAGAGAGCGAGCTTCCGGGACAAATGAACAAATCGGCC 702
 570 ----- 569
 703 TAAATTAATTAAGACATGAGATGATGCTCCGGGAACTGGGCTTCAAGACAGAGT 762
 619 GCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
 763 GCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
 679 AACATCTTGGGATTCGCGGGGACCAATATGTTTCAAGCTTATATCTGCTCCCATG 738
 823 AACATCTTGGGATTCGCGGGGACCAATATGTTTCAAGCTTATATCTGCTCCCATG 882
 739 GCGAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGTGGGATCATCCCTTAAGT 798
 883 GCGAAGGCTTATTCACAAAGCCATCATGAGAGTGGGAGTGGGATCATCCCTTAAGT 942
 799 GAGGCTCATGATTAAGAGAGTGAAGACTGCAAGTGTGCTGCAATTTCTGCTGTAAC 858
 943 GAGGCTCATGATTAAGAGAGTGAAGACTGCAAGTGTGCTGCAATTTCTGCTGTAAC 1002

859 AATGCTCAGACTCTGAGGCGCTGCTGAGTGCCTGAGGACAAACCTCCAGAGAGCTG 918
 1003 AATGCTCAGACTCTGAGGCGCTGCTGAGTGCCTGAGGACAAACCTCCAGAGAGCTG 1062
 919 CTGACCTCAGCAGAGAAACAAAGCTTTCACTGAGTGTGATGCTGCTTTCTTCT 978
 1063 CTGACCTCAGCAGAGAAACAAAGCTTTCACTGAGTGTGATGCTGCTTTCTTCT 1122
 979 AATGAGCTCTAGATCTAATGCTCAGAAAGCATTAAGCAATCTTCATCATCCGA 1038
 1123 AATGAGCTCTAGATCTAATGCTCAGAAAGCATTAAGCAATCTTCATCATCCGA 1182
 1039 GTCAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAGGCTCTGAGATCCCTCAGT 1098
 1183 GTCAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAGGCTCTGAGATCCCTCAGT 1242
 1099 GGTCTCAACAGTCCCTTGCCTCATTTGATCAAAACATCTGCAATCCGCTCAG 1158
 1243 GGTCTCAACAGTCCCTTGCCTCATTTGATCAAAACATCTGCAATCCGCTCAG 1302
 1159 TATTTGCACTTGTGGCTAATGAATATCTTCATGAACAAGCACTCCGCTGAATCCGA 1218
 1303 TATTTGCACTTGTGGCTAATGAATATCTTCATGAACAAGCACTCCGCTGAATCCGA 1362
 1219 GACAGTCTTCTGAGACTTGTGAGATGTTCTTTGCTGCTGCTGCACTGATCAAGCT 1278
 1363 GACAGTCTTCTGAGACTTGTGAGATGTTCTTTGCTGCTGCTGCACTGATCAAGCT 1422
 1279 CGATATCAACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1338
 1423 CGATATCAACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482
 1339 TGTCTTTGAAGACAGAACCGGCTTTGTCAAGCGGACCAAGCTGAGAGTCCGCTT 1398
 1483 TGTCTTTGAAGACAGAACCGGCTTTGTCAAGCGGACCAAGCTGAGAGTCCGCTT 1542
 1399 GTGCTCGGTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458
 1543 GTGCTCGGTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1602
 1459 GAGAGTACTGAGGCTGAGAGATGAATATCTGGGCTACCTTTGCTGAAACCGGAAAT 1518
 1603 GAGAGTACTGAGGCTGAGAGATGAATATCTGGGCTACCTTTGCTGAAACCGGAAAT 1662
 1519 CCTAATGGGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1578
 1663 CCTAATGGGAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1722
 1579 CTGACCTGGAATGAGGCTCGAGACAGAGACTCAAGAGACCGGGGTGAGATTTGGACC 1638
 1723 CTGACCTGGAATGAGGCTCGAGACAGAGACTCAAGAGACCGGGGTGAGATTTGGACC 1782
 1639 AGCACATATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1698
 1783 AGCACATATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842
 1699 ACTTTCCT 1746
 1843 ACTTTCCT 1890

RESULT 10
 CQ733044 1716 bp DNA linear PAT 03-FEB-2004
 LOCUS CQ733044
 DEFINITION Sequence 18978 from Patent WO02068579.
 ACCSSION CQ733044
 VERSION CQ733044.1 GI:42314969
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1

AUTHORS

Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof

JOURNAL

Patent: WO 02068579-A 18978 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES

source
1. 1716
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 81.6%; Score 1425; DB 6; Length 1716;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 84 AAAAATCGGGCTTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTCAGG 143
DB 21 AAACAAAGGCTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTCAGG 80
QY 144 CAGCAAGTCACTGTCTGGAGACCCTGCTGCTGAAAGTGTCTGGAGTCCCTT 203
DB 81 CAGCAAGTCACTGTCTGGAGACCCTGCTGCTGAAAGTGTCTGGAGTCCCTT 140
QY 204 TGTCTCTCCCGCTGGATCCCTGCTGATTAAGAACCCGACGCTGATCCCTGGGA 263
DB 141 TGTCTCTCCCGCTGGATCCCTGCTGATTAAGAACCCGACGCTGATCCCTGGGA 200
QY 264 TAACTTGGAGAGACCACTCCCTAACCCTAATTTGTCCTCCAGAACTCAGATGCTCT 323
DB 201 TAACTTGGAGAGACCACTCCCTAACCCTAATTTGTCCTCCAGAACTCAGATGCTCT 260
QY 324 CTTAAGTCAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGACCTGCT 383
DB 261 CTTAAGTCAACATGCTCAAGGTGATTAACCGAAATTCGAGTGTCAAGACCTGCT 320
QY 384 CTACCTGAACATCTATGCGCTCCGACGACGACGATCAAGGCTTCAAGCTCCCTGCT 443
DB 321 CTACCTGAACATCTATGCGCTCCGACGACGACGATCAAGGCTTCAAGCTCCCTGCT 380
QY 444 GTGTCTCCAGAGAGTGTCTCAAGCTGCTCAAGCTTCAAGCTTCAAGCTTCCGCT 503
DB 381 GTGTCTCCAGAGAGTGTCTCAAGCTGCTCAAGCTTCAAGCTTCAAGCTTCCGCT 440
QY 504 GGCTGCTTATGAGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 563
DB 441 GGCTGCTTATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 500
QY 564 CTTCACACATGAGTATGAGTATGCTCCGAGGAACTGGGCTTCAAGGACGAGTGGCTGC 623
DB 501 CTTCACACATGAGTATGAGTATGCTCCGAGGAACTGGGCTTCAAGGACGAGTGGCTGC 560
QY 624 TCTGTCTGGGTCAAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTTGTGACAT 683
DB 561 TCTGTCTGGGTCAAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTTGTGACAT 620
QY 684 CTTTGGGAGTCCGCGGAGGACCAATAGTGTCTTCAAGCTTCAAGCTTCCGAGCA 743
DB 621 CTTTGGGAGTCCGCGGAGGACCAATAGTGTCTTCAAGCTTCAAGCTTCCGAGCA 680
QY 744 AGGCTTATTCACAAACCATCATGAGAGTGGGAGTGGGAGTGGGAGTGGGAGT 803
DB 681 AGGCTTATTCACAAACCATCATGAGAGTGGGAGTGGGAGTGGGAGTGGGAGT 740
QY 804 CCAATGATTAAGAAAGTGAAGCTGCAAGTGTGTGACATTTCTGTGTGAACAATGC 863
DB 741 CCAATGATTAAGAAAGTGAAGCTGCAAGTGTGTGACATTTCTGTGTGAACAATGC 800
QY 864 GTGAGACTGTGAGGCTGTGAGGAGTGTGAGGAGCAAAACCTTCAAGGAGTGTGAC 923
DB 801 GTGAGACTGTGAGGCTGTGAGGAGTGTGAGGAGCAAAACCTTCAAGGAGTGTGAC 860

```

```

QY 924 CCTACGCCAGAAACAAAGTCTTTCATCTGAGTGTGATGAGTCTTCTTCTTAATGA 983
DB 861 CCTACGCCAGAAACAAAGTCTTTCATCTGAGTGTGATGAGTCTTCTTCTTAATGA 920
QY 984 GCGTCAAGTCTATGTGTCTGAGAAAGCAATTAAGCAATTCCTTCATCATCGAGTCAA 1043
DB 921 GCGTCAAGTCTATGTGTCTGAGAAAGCAATTAAGCAATTCCTTCATCATCGAGTCAA 980
QY 1044 TAACACAGATGTGGCTTCTGCTGCTCATGAGAGGAGCTCTGAGATCCTCAGTGGCTC 1103
DB 981 TAACACAGATGTGGCTTCTGCTGCTCATGAGAGGAGCTCTGAGATCCTCAGTGGCTC 1040
QY 1104 CAACAGTCCCTTGGCTTCATCTGATTAAGAAACATCTGACATCCCGCTCAGTATTT 1163
DB 1041 CAACAGTCCCTTGGCTTCATCTGATTAAGAAACATCTGACATCCCGCTCAGTATTT 1100
QY 1164 GCACCTTGTGCTATGATTAATCTTCATGACAAACATCCCTGACCTGAATTCAGATATTT 1223
DB 1101 GCACCTTGTGCTATGATTAATCTTCATGACAAACATCCCTGACCTGAATTCAGATATTT 1160
QY 1224 TCTTCTGAGCTTGTGAGATGTGTCTTGTGTGTCTGCACTGATCAAGCTCGATTA 1283
DB 1161 TCTTCTGAGCTTGTGAGATGTGTCTTGTGTGTCTGCACTGATCAAGCTCGATTA 1220
QY 1284 TCACAGAGATGTGTGACACCTGTCTAATCTTATGATTTCCGACCGGCTCAGTGTCT 1343
DB 1221 TCACAGAGATGTGTGACACCTGTCTAATCTTATGATTTCCGACCGGCTCAGTGTCT 1280
QY 1344 TGAAGACAGAAAGCGGCTTTTGTCAAGGCGACCAAGCTGATGAAGTCCGCTTGTGTT 1403
DB 1281 TGAAGACAGAAAGCGGCTTTTGTCAAGGCGACCAAGCTGATGAAGTCCGCTTGTGTT 1340
QY 1404 CGGTGTGCTTCTGTAAGGAGGAGCAATTTATGTTGTAAGAGCAACGAGAGAGAGAA 1463
DB 1341 CGGTGTGCTTCTGTAAGGAGGAGCAATTTATGTTGTAAGAGCAACGAGAGAGAGAA 1400
QY 1464 GTTACTGAGCGGAGATGATGAATTAATCTGAGCTTGTGCTGAAACCGGGA 1516
DB 1401 GTTACTGAGCGGAGATGATGAATTAATCTGAGCTTGTGCTGAAACCGGGA 1453

```

RESULT 11
AX958392 1857 bp DNA linear PAT 14-JAN-2004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)
Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1

Human drug metabolizing enzymes

Patent: WO 0226988-A 20 04-Apr-2002;

Incyte Genomics, Inc. (US)

Location/Qualifiers

1. 1857

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 79.4%; Score 1386.2; DB 6; Length 1857;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

```

QY 81 GGGAAAAAATGGGCTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTC 140
DB 264 GGTCAAGAGAGGCTTCTGCTGAAGGCGCAGAGAAACACAGGCTGGATGATTC 323
QY 141 GGGCAAGCAATCACTGTCTGGAGAGCTTGTCTGTGAACGTGTCTTCTGGAGTCCC 200

```

Db 324 GGGGAGAGCAAGTCACTGTCCTGGAGAGCCCTGCTGTGAAAGTGTCTCGAGAGTCCC 383
Qy 201 CTTTGTCTGCTCCCGCTGGATTCCTGTGATTTAAGAACCCGCGAGCTGATCCGCTG 260
Db 384 CTTTGTCTGCTCCCGCTGGATTCCTGTGATTTAAGAACCCGCGAGCTGATCCGCTG 443
Qy 261 GGAATTAAGTGGAGAGCAAGCTCTCAACCTTAATTTGTGCTCCAGAACTCAGAGTGGT 320
Db 444 GGAATTAAGTGGAGAGCAAGCTCTCAACCTTAATTTGTGCTCCAGAACTCAGAGTGGT 503
Qy 321 GCTCTTAAGATCAACAATGCTCAAGTGCATTACCGAAATTCGAGTGTCAAGAACTG 380
Db 504 GCTCTTAAGATCAACAATGCTCAAGTGCATTACCGAAATTCGAGTGTCAAGAACTG 563
Qy 381 CCTTAAGTCAACAATGCTCAAGTGCCTCCAGCCGATACAGAGCTCCAGCTCCGCTT 440
Db 564 CCTTAAGTCAACAATGCTCAAGTGCCTCCAGCCGATACAGAGCTCCAGCTCCGCTT 623
Qy 441 GGTGTGTGTTCCAGAGAGGCTCTCAAGAATGAGTCAAGCTCAATCTTTGATGGTCCG 500
Db 624 GGTGTGTGTTCCAGAGAGGCTCTCAAGAATGAGTCAAGCTCAATCTTTGATGGTCCG 683
Qy 501 CTTGCTGCTCAATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
Db 684 CTTGCTGCTCAATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 743
Qy 561 TTTTCTTCAACACATGAGATCAGATGCTCCGCGGAACTGCGCTTCAAGAACAGAGTGC 620
Db 744 TTTTCTTCAACACATGAGATCAGATGCTCCGCGGAACTGCGCTTCAAGAACAGAGTGC 803
Qy 621 TGTCTGTGCTGAGGCTCAAGAGAGATGAGTCTTGGTGGGAGCCCGAGCTGTGAC 680
Db 804 TGTCTGTGCTGAGGCTCAAGAGAGATGAGTCTTGGTGGGAGCCCGAGCTGTGAC 863
Qy 681 CATCTTTGGCGAGTCCGCGGAGCAATAAGTGTCTTAAGTCTTAATGCTCTCCAGTGC 740
Db 864 CATCTTTGGCGAGTCCGCGGAGCAATAAGTGTCTTAAGTCTTAATGCTCTCCAGTGC 923
Qy 741 CAAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTGACATCATCTTAACTCTGGA 800
Db 924 CAAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTGACATCATCTTAACTCTGGA 983
Qy 801 GGGCCATGATTAAG 860
Db 984 GGGCCATGATTAAG 1043
Qy 861 TGGCTCAGACTGTGAGGCTCTGAGGCTGAGGAGCAAAACCTCCAGAGAGTGCCT 920
Db 1044 TGGCTCAGACTGTGAGGCTCTGAGGCTGAGGAGCAAAACCTCCAGAGAGTGCCT 1103
Qy 921 GAGCCCTGAGCCAGAAACAAAGTCTTCACTGAGTGGTGAATGCTCTTCTCTTA 980
Db 1104 GAGCCCTGAGCCAGAAACAAAGTCTTCACTGAGTGGTGAATGCTCTTCTCTTA 1163
Qy 981 TGAAGCTCTAAGTCTAATGCTCAGAAAGCATTTAAGCAATTCCTCCATATGAGAGT 1040
Db 1164 TGAAGCTCTAAGTCTAATGCTCAGAAAGCATTTAAGCAATTCCTCCATATGAGAGT 1223
Qy 1041 CAATTAACAGAGAGTGTGCTCTGCTGCTATGAAGAGAGCTCTGAGATCTCAGAGTGG 1100
Db 1224 CAATTAACAGAGAGTGTGCTCTGCTGCTATGAGAGAGAGCTCTGAGATCTCAGAGTGG 1255
Qy 1101 CTGCAAGAGTCCCTTGGCTTCATCTGATACAAAACATCTGCAATCCGCTCAGTA 1160
Db 1256 CTGCAAGAGTCCCTTGGCTTCATCTGATACAAAACATCTGCAATCCGCTCAGTA 1274
Qy 1161 TTTTGAACCTTGTGCTAATGATTAATCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1220
Db 1275 TTTTGAACCTTGTGCTAATGATTAATCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1334
Qy 1221 CAGTCTTCTGAGACTGTGAGAGATGTGTTCTTTGTGTGCTCCGCTGAGATGAGAGAGTGC 1280

Db 1335 CAGTCTTCTGAGACTGTGAGAGATGTGTTCTTTGTGTGCTCCGCTGAGATGAGAGAGTGC 1394
Qy 1281 ATATCAAGAGATGTGCTGAG 1340
Db 1395 ATATCAAGAGATGTGCTGAG 1454
Qy 1341 CTTTGAAG 1400
Db 1455 CTTTGAAG 1514
Qy 1401 GTTGGTGTGTGCTCTCTGAG 1460
Db 1515 GTTGGTGTGTGCTCTCTGAG 1574
Qy 1461 GAGTCTTCAAG 1520
Db 1575 GAGTCTTCAAG 1634
Qy 1521 TAATGGAG 1580
Db 1635 TAATGGAG 1694
Qy 1581 GAGCTTGAACATGAGCTCTGAG 1635
Db 1695 GAGCTTGAACATGAGCTCTGAG 1749

RESULT 12
BC069548 1896 bp mRNA linear PRI 30-JUN-2004
LOCUS BC069548
DEFINITION Homo sapiens hypothetical protein FLJ31447, mRNA (CDNA clone
MGC:97193 IMAGE:7262439), complete cds.
ACCESSION BC069548
VERSION BC069548.1 GI:46854430
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1896)
REFERENCES
AUTHORS
Straussberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stepleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshilyuk, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Wortley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Kozminski, M.I., Skalske, U., Smalins, D.E.,
Schnerch, A., Schein, J.B., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
PUBMED
REFERENCES
AUTHORS
TITLE
JOURNAL
COMMENT
REMARK
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Baylor Human Genome Sequencing Center

CDNA Library Preparation: Baylor Human Genome Sequencing Center
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loubege, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRBR Plate: 4 Row: f Column: 1.

FEATURES

Source
 1..1896
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:97193 IMAGE:7262439"
 /clone_lib="PCR rescued clones"
 /clone_lib="NIH MGC 244"
 /note="Vector: PCR-Script Amp SK(+)"
 1..1896
 /gene="FLJ31547"
 /db_xref="LocusID:221223"
 120..1697
 /gene="FLJ31547"
 /codon_start=1
 /product="hypothetical protein FLJ31547"
 /protein_id="AAH69548.1"
 /db_xref="GI:46854431"
 /db_xref="LocusID:221223"
 /translation="MSGNVHDPQGLIIMVLAATPKGPSAEGRQMRRLGIMQCKO
 LVTLSPVAVNPLGVFPAAPLGSRLFTNPQASPMNLRATSPRLCLQNSKML
 LDOHMLKHYPKGVSEDCILNITAPAHADGSLPLVAFPGAFRTGSASIFDGS
 ALAAVEDVIVVVOYRLGIFGFFTTMDQAPGNMFKQVVALSVQNLISFEGDPS
 SVTIRGESAALSVSLISLSPWAKGLFHALMESGVALIPLYEADYKESDLQVAF
 FCGNNAJSDEALRLCLRTKPSKELTLLQKTSFVVVDGAFPYNEPLDLSQAFKA
 IPSIIIVNNHCEGLFPMKEAPEILISGNKSLALHIONIHLIPQVLMVANEVFD
 KHSLEIRDSLDLDLGDVFEVVPALITARYHREGATEEKEKLSRMMQVYMTAFRTD
 PNGNDLSMPAYNLTEOYLQDLNMSLQRLKEPRVDVWSTITPLILASDMLHSPLS
 SLTFLSLQPPFFPCAP"

ORIGIN

Query Match 77.0%; Score 1344.4; DB 9; Length 1896;
 Best Local Similarity 90.9%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;
 91 GGGCTTCTGCTGAAGGGCCACAGAGAAACACCGGCTGGATGATTCAAGGCAAGCA 150
 192 GGGCTTCTGCTGAAGGGCCACAGAGAAACACCGGCTGGATGATTCAAGGCAAGCA 251
 151 GTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTCCTCGAGATGCCCTTTGCTGT 210
 252 GTCACTGTGCTGGAAAGCCCTGTGCTGTGAACGTGTCCTCGAGATGCCCTTTGCTGT 311
 211 CCCCCGCTGGATATCCCTGCGATTACGAACCCGAGCCTGCATGCCCTGGAGTAACCTG 270
 312 CCCCCGCTGGATATCCCTGCGATTACGAACCCGAGCCTGCATGCCCTGGAGTAACCTG 371
 271 CGAAGAACCACTCTCACTCACTTAATTTGTGCTCGAAGCTCAAGTGGCTGCTTAAGAT 330
 372 CGAAGAACCACTCTCACTCACTTAATTTGTGCTCGAAGCTCAAGTGGCTGCTTAAGAT 431
 331 CAACACATGCTCAAGGTGCAATTACCGGAATTTGGAGTGTCAAGAGACTGCTTAACCTG 390
 432 CAACACATGCTCAAGGTGCAATTACCGGAATTTGGAGTGTCAAGAGACTGCTTAACCTG 491
 391 AACATCTATGCGCTGCGCCACGCGGATACAGGCTCCCAAGCTCCCGCTTGGTGTGCTTC 450
 492 AACATCTATGCGCTGCGCCACGCGGATACAGGCTCCCAAGCTCCCGCTTGGTGTGCTTC 551

QY 451 CCAGAGGTGCTTCAAGACTGAGCTCAGCTCCATCTTTGATGGATCCGCGCTGCTGCC 510
 DB 552 CCAGAGGTGCTTCAAGACTGAGCTCAGCTCCATCTTTGATGGATCCGCGCTGCTGCC 611
 QY 511 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
 DB 612 TATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
 QY 571 ACATGGAGTCAAGATGCTCCGCGGAACTGGGCTTCAAGAGCCAGAGTGGCTGCTGCTGCT 630
 DB 672 ACATGGAGTCAAGATGCTCCGCGGAACTGGGCTTCAAGAGCCAGAGTGGCTGCTGCTGCT 731
 QY 631 TGGGTCCAGAAAGATCATGAGTCTTCCGTGGGAGCCCACTGCTGTAACCATCTTGGC 690
 DB 732 TGGGTCCAGAAAGATCATGAGTCTTCCGTGGGAGCCCACTGCTGTAACCATCTTGGC 791
 QY 691 GAGTCCGGGGAGCATAGTGTCTTAAGTCTTAAGTCTTCCATGAGCCAAAGCTTA 750
 DB 792 GAGTCCGGGGAGCATAGTGTCTTAAGTCTTAACTGCTTCCATGAGCCAAAGCTTA 851
 QY 751 TTCCCAAAAGCATATGAGAGAGTGGGAGCATATCCCTTACCTGAGAGCCCATGAT 810
 DB 852 TTCCCAAAAGCATATGAGAGAGTGGGAGCATATCCCTTACCTGAGAGCCCATGAT 911
 QY 811 TATGAGAGAGTGAAGACCTGACAGTGTGACATTTCTGTGTAA CAA TGCCTCAGAC 870
 DB 912 TATGAGAGAGTGAAGACCTGACAGTGTGACATTTCTGTGTAA CAA TGCCTCAGAC 971
 QY 871 TCTGAGGCGCTGCTGAGAGTGTGCTGAGGACAAACCCCTCAAGAGAGCTGACCTTCAGC 930
 DB 972 TCTGAGGCGCTGCTGAGAGTGTGCTGAGGACAAACCCCTCAAGAGAGCTGACCTTCAGC 1031
 QY 931 CAGAAACAAAGCTTCTCACTGAGTGTGATGTGCTTTCTTCTATGAGCTCTA 990
 DB 1032 CAGAAACAAAGCTTCTCACTGAGTGTGATGTGCTTTCTTCTATGAGCTCTA 1091
 QY 991 GATCTATTTCTGAGAAAGATTTAAAGCAATTCCTTCACTATGAGAGTCAATTAACAC 1050
 DB 1092 GATCTATTTCTGAGAAAGATTTAAAGCAATTCCTTCACTATGAGAGTCAATTAACAC 1151
 QY 1051 GAGTGTGCTTCTGCTGCTGCTTATGAGAGAGCTCCTGAGATCCTCAGTGGCTCAACAG 1110
 DB 1152 GAGTGTGCTTCTGCTGCTGCTTATGAGAGAGCTCCTGAGATCCTCAGTGGCTCAACAG 1211
 QY 1111 TCCCTTGCCTTCACTGATTAACAAACATCTGACATCCGCTCAGTATTTGACCTT 1170
 DB 1212 TCCCTTGCCTTCACTGATTAACAAACATCTGACATCCGCTCAGTATTTGACCTT 1271
 QY 1171 GTGGCTAATGAATATCTTCATGACAGCACTCCCTGACCTGAATCCGAGACAGTCTTCTG 1230
 DB 1272 GTGGCTAATGAATATCTTCATGACAGCACTCCCTGACCTGAATCCGAGACAGTCTTCTG 1331
 QY 1231 GACTTGTGAGATGTGTTCTTTGTGCTGCTGACCTGATCAAGCTCGATATCAGAGA 1290
 DB 1332 GACTTGTGAGATGTGTTCTTTGTGCTGCTGACCTGATCAAGCTCGATATCAGAGA 1391
 QY 1291 GATGTGTGACACTGTGTACTTATATGATTTGGGACCGGCTCAGTGGCTTTGAAGAC 1350
 DB 1392 ----- 1391
 QY 1351 ACGAAGCGGCTTTTGTCAAGCCGACGACTGATGAAGTCCGTTGTGCTGGTGGT 1410
 DB 1392 ----- 1391
 QY 1411 GCTTCTGAGAGGGGAGATGTTATGTTCAGAGAGGCGACGAGAGAGAGATTAAGT 1470
 DB 1392 -----GAAAGGCGAAGGAGGAGGAGAGATTAAGT 1421
 QY 1471 AGCCGAAAGATGAGAAATTAAGTGGCTTCTGCAACCGGAAATCTTAATGGGAAC 1530
 DB 1422 AGCCGAAAGATGAGAAATTAAGTGGCTTCTGCAACCGGAAATCTTAATGGGAAC 1481
 QY 1531 GACCTGTCTGTGCGCCAGCTTAATCTGACTGAGAGATCACTCAGCTGAGCTTGAAC 1590

|||||
Db 1482 GACCTGTCTCTGTGGCCAGCTTATATCTGACTGAGCAGTACTCTCAGCTGGACCTTGAAC 1541
Qy 1501 ATGAGCCTTCGACAGAGACTCAAAGAACCGCGGGTGGATTTTGGACACGACATCCCC 1650
Db 1542 ATGAGCCTTCGACAGAGACTCAAAGAACCGCGGGTGGATTTTGGACACGACATCCCC 1601
Qy 1651 CTGATCTGTCTGCGCTCCGACAGTCCGACAGTCCCTTCTTCTTAACCTTCCCTCT 1710
Db 1602 CTGATCTGTCTGCGCTCCGACAGTCCGACAGTCCCTTCTTCTTCTTAACCTTCCCTCT 1661
Qy 1711 CTCTTCAGAGCTTTCTTTTCTTTTGTGCTCCCTGA 1746
Db 1662 CTCCTCAAGCCTTTCTTTTCTTTTGTGCTCCCTGA 1697

RESULT 13
AX714052 1962 bp DNA linear PAT 15-APR-2003
LOCUS Sequence 736 from Patent EP1293569.
ACCESSION AX714052
VERSION AX714052.1 GI:29888980
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Iwagui,T., Sugiyama,T., Otsuki,T., Makamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Maehara,Y.
TITLE Full-length cDNA
JOURNAL Patent: EP 1293569-A 736 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)
FEATURES
Source 1. 1962
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 77.0%; Score 1344.4; DB 6; Length 1962;
Best Local Similarity 90.9%; Pred. No. 0;
Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

Qy 91 GGGCCTTCTGCTGAGGGCCACAGAGAACACAGGCTGGATTCAGGGCAAGCA 150
Db 222 GGGCCTTCTGCTGAGGGCCACAGAGAACACAGGCTGGATTCAGGGCAAGCA 281
Qy 151 GTCACTGTGTGGAGAGCCCTGTGCTGTGAAGCTGTTCCTGAGAGTCCCTTGTCTGT 210
Db 282 GTCACTGTGTGGAGAGCCCTGTGCTGTGAAGCTGTTCCTGAGAGTCCCTTGTCTGT 341
Qy 211 CCCCCTGGGATCCCTGCGATTTACGAACCCGCGAGCTGATCCGCTGGGATTAATTG 270
Db 342 CCCCCTGGGATCCCTGCGATTTACGAACCCGCGAGCTGATCCGCTGGGATTAATTG 401
Qy 271 CGAGAGGCACTCTTACCTTAATTGTGCTCCAGAACTCAGAGTGGCTGCTTAAGT 330
Db 402 CGAGAGGCACTCTTACCTTAATTGTGCTCCAGAACTCAGAGTGGCTGCTTAAGT 461
Qy 331 CAACATGCTCAAGGTGATTAACCGAAATTCGAGGTGCGAAGACTGCTTAACTG 390
Db 462 CAACATGCTCAAGGTGATTAACCGAAATTCGAGGTGCGAAGACTGCTTAACTG 521
Qy 391 AACATCTATGCGCCCGCCAGCGCGATACAGGCTCAAGGCTCCGCTTGGTGTGCTTC 450
Db 522 AACATCTATGCGCCCGCCAGCGCGATACAGGCTCAAGGCTCCGCTTGGTGTGCTTC 581
Qy 451 CCAGAGGTGCTTCAAGACTGCTCAGGCTCATCTTGTGATGGTCCGCTGGCTGCC 510
|||||

Db 582 CCAGAGGTGCTTCAAGACTGCTCAGGCTCATCTTGTGATGGTCCGCTGGCTGCC 641
Qy 511 TATGAGAGCTGCTGGTGTGTGCTGCTAGTACGGGCTTAGGAATATTTGGTTCTTCAAC 570
Db 642 TATGAGAGCTGCTGGTGTGTGCTGCTAGTACGGGCTTAGGAATATTTGGTTCTTCAAC 701
Qy 571 ACATGGGATCAGACTGCTCCGCGGAACTGGGCTTCAAGAACCAAGTGGCTGCTGTGCC 630
Db 702 ACATGGGATCAGACTGCTCCGCGGAACTGGGCTTCAAGAACCAAGTGGCTGCTGTGCC 761
Qy 631 TGGGTCCAGAGACATGAGTCTTCTGGTGGGAAACCCAGCTCTGTGTGACATCTTTGGC 690
Db 762 TGGGTCCAGAGACATGAGTCTTCTGGTGGGAAACCCAGCTCTGTGTGACATCTTTGGC 821
Qy 691 GAGTCCGCGGAGCCATAGTGTCTTCTAGTCTTAATCTGTCCATGGCCAAAGGCTTA 750
Db 822 GAGTCCGCGGAGCCATAGTGTCTTCTAGTCTTAATCTGTCCATGGCCAAAGGCTTA 881
Qy 751 TTCACAAAGCCATGATGAGAGAGTGGGTGGCCATATCCCTTACCTGGAGCCCATGAT 810
Db 882 TTCACAAAGCCATGATGAGAGAGTGGGTGGCCATATCCCTTACCTGGAGCCCATGAT 941
Qy 811 TATGAGAGAGTGAAGACCTGAGGTGTGTCACATTTCTGTGTAACTATCCGTCAAG 870
Db 942 TATGAGAGAGTGAAGACCTGAGGTGTGTCACATTTCTGTGTAACTATCCGTCAAG 1001
Qy 871 TCTGAGGCCCTCTGAGGTGCTGAGCAAAACCTTCCAGAGACTGCTGACCTCAGC 930
Db 1002 TCTGAGGCCCTCTGAGGTGCTGAGCAAAACCTTCCAGAGACTGCTGACCTCAGC 1061
Qy 931 CAGAAAACAAAGTCTTCACTCGAGTGTGATGTGCTTCTTCTTAATAGGCTCTA 990
Db 1062 CAGAAAACAAAGTCTTCACTCGAGTGTGATGTGCTTCTTCTTAATAGGCTCTA 1121
Qy 991 GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTTCCATCATGTGAGTCAATTAACAC 1050
Db 1122 GATCTATTGTCTCAGAAAGCATTTAAAGCAATCTTCCATCATGTGAGTCAATTAACAC 1181
Qy 1051 GAGTGTGGCTTCTGCTGCTGCTATGAGAGAGGCTCCTGAGATCCTGAGTGGCTCAACAG 1110
Db 1182 GAGTGTGGCTTCTGCTGCTGCTATGAGAGAGGCTCCTGAGATCCTGAGTGGCTCAACAG 1241
Qy 1111 TCCCTTGGCCCTCATCTGATACAAACATCCTGACATCCGCTCAGTATTTGACCTT 1170
Db 1242 TCCCTTGGCCCTCATCTGATACAAACATCCTGACATCCGCTCAGTATTTGACCTT 1301
Qy 1171 GTGGCTAATGAATTAATTCATGACAAAGACTCCTGACTGAATTCGAGACAGTCTTCTG 1230
Db 1302 GTGGCTAATGAATTAATTCATGACAAAGACTCCTGACTGAATTCGAGACAGTCTTCTG 1361
Qy 1231 GACTTGTGAGATGTGTTCTTTGTGTGCTCCTGACATGATCAAGCTTCGATATACAGA 1290
Db 1362 GACTTGTGAGATGTGTTCTTTGTGTGCTCCTGACATGATCAAGCTTCGATATACAGA 1421
Qy 1291 GATGTGGTGACCTGTACTTATAGAGTTTGGGACCCGCTCAAGTCTTTGAAGAC 1350
Db 1422 ----- 1421
Qy 1351 ACGAAGCCGCTTTTGTCAAAAGCCAGCAGCTGATGAAGTCCGTTGTGTGCTGTGT 1410
Db 1422 ----- 1421
Qy 1411 GCCTTCTGAAGGGGAGCAATTGTATGTTCAGAAAGCCACGAGAGAGAAAGTTACTG 1470
Db 1422 -----GAAAGAGCCACGAGAGAGAAAGTTACTG 1451
Qy 1471 AGCGGAAGATGATGAATACTGGGCTACTTTGTCTGAAACCGGGAATCTTAATGGGAAC 1530
Db 1452 AGCGGAAGATGATGAATACTGGGCTACTTTGTCTGAAACCGGGAATCTTAATGGGAAC 1511
Qy 1531 GACCTGTCTCTGTGGCAGCTTATATCTGAATCTGAAGAGTACCTCAGCTGGACTTGAAC 1590
Db 1512 GACCTGTCTCTGTGGCAGCTTATATCTGAATCTGAAGAGTACCTCAGCTGGACTTGAAC 1571
|||||

Qy 1591 ATGAGCTCGGAGAGAGACTCAAGAAACCGCGGCTGATTTTGGACCAACCATCTCCC 1650
 Db 1572 ATGAGCTTCGAGACAGAGACTCAAGAAACCGCGGCTGATTTTGGACCAACCATCTCCC 1631
 Qy 1651 CTGATCTGTCTGCTCGCTCGAATGCTCCACAGTCTCTTTCTTCTTAACCTTCTCTCT 1710
 Db 1632 CTGATCTGTCTGCTCGCTCGAATGCTCCACAGTCTCTTTCTTCTTAACCTTCTCTCT 1691
 Qy 1711 CTCTCGAGCTTCTTTCTTTCTTTGCTCTTGA 1746
 Db 1692 CTCTCGAGCTTCTTTCTTTCTTTGCTCTTGA 1727

RESULT 14
 AK056109 1962 bp mRNA linear PRI 30-JAN-2004
 LOCUS Homo sapiens cDNA FLJ1547 fib, clone NT2R12001010, weakly similar
 DEFINITION to FATTY ACYL-COA HYDROLASE PRECURSOR, MEDIUM CHAIN (EC 3.1.2.14).
 ACCESSION AK056109
 VERSION AK056109.1 GI:16551422
 KEYWORDS oligo capping; fib (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
 Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
 Sekine, M., Odayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
 Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
 Nagahari, K., Murakami, K., Yasuda, K., Iwayanagi, T., Wagatsuma, M.,
 Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
 Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kondo, H.,
 Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
 Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
 Hirooka, K., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
 Hiroda, M., Hikota, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,
 Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
 Takeuchi, K., Arita, M., Imose, N., Mueshino, K., Yuuki, H., Oshima, A.,
 Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
 Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
 Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
 Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
 Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
 Kunagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
 Tachibana, H., Tanigami, A., Fujizawa, T., Ono, T., Yamada, K., Fujii, Y.,
 Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
 Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,
 Nakajima, Y., Itoh, T., Shigeta, K., Senba, T., Matsunuma, K.,
 Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
 Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
 Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R.,
 Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T., and Sugano, S.
 Complete sequencing and characterization of 21,243 full-length
 human cDNAs
 Nac. Genet. 36 (1), 40-45 (2004)
 14702039

TITLE
 JOURNAL PUBMED
 REFERENCE
 AUTHORS
 Nishimura, K., Magatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
 Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, B., Omura, Y., Abe, K.,
 Komihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,
 Kikkawa, E., Yamashita, H., Matsuno, K., Nakamura, Y., Sekine, M.,
 Kikuchi, H., Murakami, K., Kanehori, K., Takahashi, Fujii, A.,
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahari, K., Masuno, Y., Nagai, K., and Isegai, T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 1962)

AUTHORS

Isegai, T., Otsuki, T. and Sugiyama, T.
 Direct Submission
 Submitted (24-Oct-2001) Takao Isegai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yano, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp; Tel: 81-438-52-3975; Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: RAB and
 HRI.

COMMENT

FEATURES

source

Location/Qualifiers
 1..1962
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="NT2R12001010"
 /cell_line="NT2"
 /cell_type="teratocarcinoma"
 /clone_id="NT2R12"
 /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
 precursor cells treated 2-weeks mitotic inhibitor after
 5-weeks retinoic acid (RA) induction.-majority NT2 neuron"
 150..1727
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB71094.1"
 /db_xref="GI:16551423"
 /translation="MSGNWTVPQGIILMAWTVAAPTKPSANGPQNRRLIGIQKQ
 VTUGSPVPPNVLGTPFAAPPLGSLFNTNPQASPMDLREXTSPNLCLOSEML
 LDHMLKVRPKGVSDCLYLNIVAPHADGSKLIVWPEFGAKFGKSAISIPGSS
 ALAAVEBLVAVVQYRIGIEFPTWDQAHGMAFKDQVLAALSHYKESIDQVVAH
 SVTFPESGALISVSSILISPAKGLFHKRIMSGVAILPYEAHDEKESIDQVVAH
 FCNNASDSBALRCLRTKSKELTLTSQKTSFTRVVGAFPPNPDLDSQAKKA
 ISFIIGNNHEGCEFLPMKEAPRLLSGNSKSLALHILQNIHNPPIYLHVAHYRFD
 KSLTLRPSLDLIDGDPVPPVPAITTAHREHGAETBEKLSRKKMKWATFARGN
 PNDNSLMPFAVLTQYVQLDLNLSLQRLKPEPRVPTSTPLILSASDMLHSPLS
 SLTFSLDPPFFFCAP"

CDS

ORIGIN

Query Match 77.0%; Score 1344.4; DB 9; Length 1962;
 Best Local Similarity 90.9%; Matched No. 0;
 Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

Qy 91 GGGCTTTCTGCTGAGAGGCTCAAGAGAAACCAAGGCTGGATGATTCAGGCAAGCA 150
 Db 222 GGGCTTTCTGCTGAGAGGCTCAAGAGAAACCAAGGCTGGATGATTCAGGCAAGCA 281
 Qy 151 GTCACGTGTGTGGAGAGCCCTGTGCTGGAAGCTTCTCTGAGATCCCTTTGCTGCT 210
 Db 282 GTACATGTGTGGAGAGCCCTGTGCTGGAAGCTTCTCTGAGATCCCTTTGCTGCT 341
 Qy 211 CCCCCTGTGGAGATCCCTGTGCTGGAAGATCCGACGCTGCATCGCCCTGGATTAATTG 270
 Db 342 CCCCCTGTGGAGATCCCTGTGCTGGAAGATCCGACGCTGCATCGCCCTGGATTAATTG 401
 Qy 271 CGAAGAGCCACCTCTCTCAACCTTAATTTGTGCTGCTCAAGAACTCAAGTGTGCTTTAGAT 330
 Db 402 CGAAGAGCCACCTCTCTCAACCTTAATTTGTGCTGCTCAAGAACTCAAGTGTGCTTTAGAT 461
 Qy 331 CAACATGCTCAAGATGATTAACCGAATTCGAGTGTGAGAGACATGCTCACTG 390
 Db 462 CAACATGCTCAAGATGATTAACCGAATTCGAGTGTGAGAGACATGCTCACTG 521
 Qy 391 AACATGTATGCTGCTGCCACGCGCATCAAGGCTCAAGCTCCCGCTTGTGTGTTTC 450
 Db 522 AACATGTATGCTGCTGCCACGCGCATCAAGGCTCAAGCTCCCGCTTGTGTGTTTC 581
 Qy 451 CGAAGAGGCTTCAAGATGCTCAAGCTTCACTTTGATGGTCCGCGCTGGCTGCC 510
 Db 582 CGAAGAGGCTTCAAGATGCTCAAGCTTCACTTTGATGGTCCGCGCTGGCTGCC 641

Db 133 GCACTGTGCTGGAGAGACCATGCTGTGAACGTGTTCTCTGGAGATCCCTTGTGCG 192
Qy 211 CCCCCGCTGGAGATCCCTGCGATTTACGAACCCGAGCTGCATGCCCTGGATTAACCTTG 270
Db 193 CCCCCCTTAAGAACCCCTGCGATTTAAGAGACCCAAAGCTGCATGCTCGGAATGACTCA 252
Qy 271 CGAAGAGCACCTCTCAACCTTAATTTGGCTCCAGAACTCAGAGTGGCTGCTTATGAT 330
Db 253 CGAATGCGACAGCTCTAATCTTAATTTGGCTCCAGAACTCAGTGTGGCTGCTCAAGT 312
Qy 331 CAACAATGCTCAAGTGCATTTACCGAATTCGAGTGTCAAGAACTGCTCAACCTG 390
Db 313 CAACAATTTCTCAAGTGCATTTACCGAATTTAGAAATGTCCGAAGCTGCTCAACCTT 372
Qy 391 AACATATGAGGCTGCGCAACCGGATACAGGCTCCAAAGTCCCGCTTGTGTGGTTC 450
Db 373 AACATCTACGACCAAGCCGCAACCGGCTCCAAAGCTCCCGTAATGTGTGGTTC 432
Qy 451 CCAAGAGGTGCTTCAAGACTGGCTCAGCTCCATCTTTATGAGGTCGCGCTGGCTGCC 510
Db 433 CTTGAGAGTGCCTTTGAGACTGGCTCAGCTCCATCTTGAATGGTTCGCGCTGGCTGCC 492
Qy 511 TATGAGAGCTGCTGCTGT 570
Db 493 TATGAGAGT 552
Qy 571 ACATGGAGTACAGATGCTCCGAGGAACTGGAGCTTCAAGAGACAGGCTGCTGTGTGC 630
Db 553 ACAGAGGACAGAGACGCTCCGAGGAACTGGAGCTTCTGAGACAGTGTGTGTGTGTGTGT 612
Qy 631 TGGGTCAAGAAACATCGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 690
Db 613 TGGGTCAAGAAACATCGAGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 672
Qy 691 GAGTCCGAGGAGCCATAAGT 750
Db 673 GAGTCCGAGGAGCCATAAGT 732
Qy 751 TTCCCAAGAGCCATCATGAGAGTGGAGTGGCCATCATCTTACCTGAGAGCCCATGAT 810
Db 733 TTCCCAAGAGCCATCATGAGAGTGGAGTGGCCATCATCTTACCTGAGAGCCCATGAT 792
Qy 811 TATGAGAGAGTGAAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 870
Db 793 GATGAGAGAGTGAAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 852
Qy 871 TCTGAGGCTGCTGAGTGTGCTGAGCAAAACCTCCAAAGAGCTGAGCCCTGAGC 930
Db 853 TCCGTAGCTGTGCTCAAGTGTGCTGAGGCAAAATCTCCGAGAGTGTGAGACATCAAC 912
Qy 931 CAGAAAACAAAGTCTTTCACTCGAGTGTGATGTGCTTTCTTTCTTATGAGCTCTTA 990
Db 913 AAGAAAACAAAGTCTTTCACTCGAGTGTGATGTGCTTTCTTTCTTATGAGCTCTTA 972
Qy 991 GATCATTTGTCTCAGAAAGCATTTAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1050
Db 973 GACCTATTGACTGAGAAAACATTTAATTCATTCCTTCTGATCGAGTCAATTAACAC 1032
Qy 1051 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
Db 1033 GAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
Qy 1111 TCCCTTGCCTCATCTGATCAAAACATCTCGACATCCCGCTCGATATTGACCTTT 1170
Db 1093 TCTCTGCTCCCTGACCTTGTATCCAGAGTCTGACATCCCAACAGTACTTATATCTT 1152
Qy 1171 GTGCTTATGATATCTTTCATGACAAAGACATCCCTGACTGAAATCCGAGACAGTCTTCTG 1230
Db 1153 GTGCTTATGATATCTTTCATGACAAAGACATCCCTGACTGAAATCCGAGTACTTATCTG 1212
Qy 1231 GACTTGTGAGAGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1290

Db 1213 GACTTGTGAGAGTGTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1272
Qy 1291 GATGCTGAGAGACCTGTCTAATCTTATGATTTGGGACACCGGCTCAAGTCTTGAAGAC 1350
Db 1273 GATGCTGAGAGACCTGTCTAATCTTATGATTTGGGACACCGGCTCAAGTCTTGAAGAC 1332
Qy 1351 ACGAAGCCGCTTTTGTCAAAAGCCGACAGCTGTATGAAGTCCGCTTGTGTGTGTGTGT 1410
Db 1333 ACGAAGCCGCTTTTGTCAAAAGCCGACAGCTGTATGAAGTCCGCTTGTGTGTGTGTGT 1392
Qy 1411 GCTTCTGAAAGGAGGACATGTATTGTTTCAAGAGCCACGAGAGAGAGAAATTTGCTG 1470
Db 1393 GCTTCTGAAAGGAGGACATGTATTGTTTCAAGAGCCACAGAGAGAGAAATTTGCTG 1452
Qy 1471 AGCCGAGAGATGATGAATTAATCTGGGCTACTTGTGTGAAACCGGGAATCTTAATGGAAAC 1530
Db 1453 AGCAGAGAGATGATGAATTAATCTGGGCTACTTGTGTGAAACCGGGAATCTTAATGGAAAC 1512
Qy 1531 GACTGTCTGTGTGACAGCTTATATCTGACTGAGACATCTCCAGCTGACTTGAAC 1590
Db 1513 GACTGTCTGTGTGACAGCTTATATCTGACTGAGACATCTCCAGCTGACTTGAAC 1572
Qy 1591 ATGAGCTCTGAGACAGACTGAAGAAACCGGAGTGAATTTTGAACACAGACCATCCC 1650
Db 1573 ATGAGCTCTGAGACAGACTGAAGAAACCGGAGTGAATTTTGAACACAGACCATCCC 1632
Qy 1651 CTGATCTGTGTGCTTCCGACATGCTCCAGAGTCTTCTTCTTAACTTCTCTCT 1710
Db 1633 CTGATTAATGTCATGCTCCAGACACCCCGTGTCTCTCTCTTAACTTCTCTCTCT 1692
Qy 1711 CTCTCCAGCTTCTTCTTCTTCTTGTGTCTTGA 1746
Db 1693 GTGCTCTGCTTCTTATCTCTCTGCTCTTGA 1728

Search completed: June 14, 2005, 22:52:51
Job time : 7589.61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 5793.46 Seconds
(without alignments)
11471.585 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 1746

Sequence: 1 atgcacacagagactacttc.....ttttctttgtgcctcttga 1746

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	479	27.4	742	4	BI822069 603039938
2	418.6	24.0	1680	9	AY414461 Mus muscu
3	414.6	23.7	576	4	BI824830 603033738
4	406.6	23.3	1680	9	AY414459 Homo sapi
5	397.8	22.8	2661	3	AK033563 Mus muscu
6	396.2	22.7	2687	3	AK077248 Mus muscu
7	387.6	22.2	4927	3	AK040349 Mus muscu
8	383.6	22.0	1958	3	AK078953 Mus muscu
9	355.2	20.3	1944	3	AK078879 Mus muscu
10	345.8	19.8	2888	3	HSB803013 Homo sapi
11	344.2	19.7	3909	3	HSB806270 Homo sapi
12	338	19.4	1918	7	CP111083 Shultzom
13	331.4	19.0	2038	3	BC019926 Mus muscu
14	298.4	17.1	1826	3	CR618303 Ttll-1eng
15	293.6	16.8	1859	3	CR641368 Tetradon
16	292	16.7	1817	3	CR634885 Tetradon
17	289	16.6	1610	3	AK050444 Mus muscu
18	286	16.4	2356	3	AK037191 Mus muscu
19	285	16.3	1491	5	BX280446 BX280446
20	284.2	16.3	1849	3	CR647187 Tetradon
21	284	16.3	539	5	BUS85267 W1701766V
22	282.2	16.2	805	5	AI808985 AF67405.X
23	282.2	16.2	808	5	BP442124 BP442124
24	282.2	16.2	808	5	BP446194 BP446194

25	282.2	16.2	840	5	BP445396	BP445396
26	282.2	16.2	841	5	BP447764	BP447764
27	282.2	16.2	860	5	BP441209	BP441209
28	282	16.2	904	5	BO900930	AGENCOURT
29	281.8	16.1	798	5	BP444116	BP444116
30	281.8	16.1	808	5	BP443157	BP443157
31	281.8	16.1	877	5	BP441564	BP441564
32	281.8	16.1	877	5	BP442855	BP442855
33	281.6	16.1	1823	3	CR640788	CR640788
34	280.8	16.1	1853	3	CR723405	CR723405
35	280.2	16.0	840	5	BP442111	BP442111
36	279.8	16.0	783	5	BP445505	BP445505
37	279.6	16.0	792	5	BP443197	BP443197
38	276	15.8	645	5	BUS85270	BUS85270
39	275.4	15.8	905	7	CO583389	ILUWIGEN
40	274.8	15.7	989	3	AK007235	AK007235
41	274	15.7	1857	3	CR637401	CR637401
42	273.6	15.7	1117	7	CR61747	CR61747
43	272.4	15.6	763	7	CR765061	CR765061
44	272	15.6	770	5	BP444251	BP444251
45	271.6	15.6	906	5	BUS49807	BUS49807

ALIGNMENTS

RESULT 1
BI822069
LOCUS 603039938P1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5181056 5', mRNA sequence.
DEFINITION BI822069
ACCESSION BI822069.1 GI:15933619
VERSION BI822069.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 742)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: L14M11451 row: k column: 09
High quality sequence stop: 740.
Location/Qualifiers
1..742
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="IMAGE:5181056"
/lab_host="DH10B"
/clone_1lb="NIH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 27.4%; Score 479; DB 4; Length 742;
Best Local Similarity 100.0%; Pred. No. 1.4e-125;
Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 GGGCTTCTGCTGAAGGCGCACAGAGAACACCAAGCTGGATGGATTCAAGGGCAAGCAA 150
DB 103 GGGCTTCTGCTGAAGGCGCACAGAGAACACCAAGCTGGATGGATTCAAGGGCAAGCAA 162
QY 151 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGCTTCCCGAGTCCCTTTGCTGCT 210
DB 163 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGCTTCCCGAGTCCCTTTGCTGCT 222
QY 211 CCCCCGCTGGAGATCCCTGCGATTACAGAACCCGAGCTGCATGCCCTGGATTAATCTTG 270
DB 223 CCCCCGCTGGAGATCCCTGCGATTACAGAACCCGAGCTGCATGCCCTGGATTAATCTTG 282
QY 271 CGAAGAGCGACCTCTCAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTTTAGAT 330
DB 283 CGAAGAGCGACCTCTCAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTTTAGAT 342
QY 331 CAACACATGCTCAAGGTGCAATTACCCGAAATTCGAGTGTCAAGAGCTGCTTAAGAT 390
DB 343 CAACACATGCTCAAGGTGCAATTACCCGAAATTCGAGTGTCAAGAGCTGCTTAAGAT 402
QY 391 AAGATCATGCGCTGCTCCACCGCGATACAGGCTCCAGCTCCCTGTGTGTGCTG 450
DB 403 AAGATCATGCGCTGCTCCACCGCGATACAGGCTCCAGCTCCCTGTGTGTGCTG 462
QY 451 CGAAGAGTGCCTTCAAGACTGTGCTCAGCTCCATCTTGAATGGTCCCGCTGCTGC 510
DB 463 CGAAGAGTGCCTTCAAGACTGTGCTCAGCTCCATCTTGAATGGTCCCGCTGCTGC 522
QY 511 TATGAGAGCTGCTGCTGTGTGTGCTGCTCAGTACCGGCTAGAAATATTGTTCTTAC 569
DB 523 TATGAGAGCTGCTGCTGTGTGTGCTGCTCAGTACCGGCTAGAAATATTGTTCTTAC 581

RESULT 2
AY414461 1680 bp DNA linear GSS 17-DEC-2003
LOCUS AY414461
DEFINITION Mus musculus CES2 gene, VIRUTAL TRANSCRIPT, partial sequence.
ACCESSION AY414461
VERSION AY414461.1 GI:39770423
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1680)
Clark,A.G., Gianoweki,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanendauum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Interfering nonneutral evolution from human-chimp-mouse orthologous
gene crie
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1680)
AUTHORS Clark,A.G., Gianoweki,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanendauum,D.M., Civeello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1680
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

gene <1..>1680
/gene="CES2"
/locus_tag="HCM5231"

ORIGIN

Query Match 24.0%; Score 418.6; DB 9; Length 1680;
Best Local Similarity 56.9%; Pred. No. 3.8e-108;
Matches 885; Conservative 0; Mismatches 634; Indels 36; Gaps 5;

QY 115 AGAACAACCAAGGCTGGAGATGATTCAGGGCAAGCAAGTCACTGTCTGGAAAGCCCTGTG 174
DB 100 AGAACAACCAAGGCTGGAGATGATTCAGGGCAAGCAAGTCACTGTCTGGAAAGCCCTGTG 159
QY 175 CCTGGAAGCTGTCTGAGATCCCTTTGCTGCTGCCCGCTGGAGATCCCTGGATTT 234
DB 160 GGTGTCACTCTTCTCTGGAGATCCCTTTGCTGCTGCCCGCTGGAGATCCCTGGATTT 219
QY 235 AGAACCCTGAGCTGCATCCGCTGGAGATTAATCTTGAGAGAGCACTCTTACCTTAAT 294
DB 220 GCAACCCCTGAGCACTGAAACCAATGGGCTGTGTGAGAGATGGAGATCAATCCGCGC 279
QY 295 TTGTGCTTCAG---AATCAGAGTGGCTGCTTAATGATCAACATGCTCAAGTGCAT 351
DB 280 ATGTGTGAGAGATCAATGCAATGAATATGAGAGCTTTTAAGCTGTGAAGCTAAC 339
QY 352 TACCGAAATTCGAGATGCTCAAGAACTGCTTACCTGAACATCTAATGGGCTGCCAC 411
DB 340 TTGCTCTCCCTTCCCTAATGCTGAGAGCTGCTTAATCAATCTAATGAGCAAGATCAT 399
QY 412 GCCGATACAGGCTCCAGAGCTCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471
DB 400 GCCGATACAGGCTCCAGAGCTCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 459
QY 472 GGTTCAGCTTCATCTTGAATGGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
DB 460 GGTTCAGCTTCATCTTGAATGGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 532 GTGTTCAGTACCGGCTGAGAAATATTGTTCTTCAACATGAGATCAGATGCTCCG 591
DB 520 ACTATTCAGTACCGGCTGAGAAATATTGTTCTTCAACATGAGATCAGATGCTCCG 579
QY 592 GGAATCTGAGCTTCAAGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
DB 580 GGAATCTGAGCTTCAAGAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639
QY 652 TTCTTGTGTGGAGACCCCAAGCTGTGTGACATCTTGTGAGATCCCGGAGGCAATAGT 711
DB 640 CACTTGTGTGGAGACCCCAAGCTGTGTGACATCTTGTGAGATCCCGGAGGCAATAGT 699
QY 712 GTTCTAGCTTAACTGATCTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAG 771
DB 700 GTTCTAGCTTAACTGATCTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAG 759
QY 772 AGTGGGTGGCCATCATCTTACCTGAGGCGCAATGATATGAGAGAGTGAAGACCTG 831
DB 760 AGTGGGTGGCCATCTGCTGCTTACCTTATCTCAACACTCTGAGA-----TGCTCC 813
QY 832 CAGGTGTGACATTTCTGTGTAACTATGCTGAGACTGTGAGGCTCTGTGAGGCTG 891
DB 814 ACTACGTGTGCAAGCTCTGTGATGTAGAGGCTGTGAGCTGAGAGGCTGTGTGAGGCTG 873
QY 892 CTGAGAGCAAAACCTTCAAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTG 951
DB 874 CTGAGAGCAAAAGTGAAGAGAGAGATTTGCTATTAAGAGAGCTTCAAGATGATTCCT 933
QY 952 CAGGTGTGATGTGTGCTTTCTTCTTAATGAGCTTGAATCTAATGTCTCAGAAAGCA 1011
DB 934 GGT 993
QY 1012 TTTAAGCAATTCCTTCATCATGAGATCAATTAACAGAGTGTGCTTCTGCTGCT 1071
DB 994 TTTAAGCAATTCCTTCATCATGAGATCAATTAACAGAGTGTGCTTCTGCTGCT 1053

QY 1072 -----ATGAGAGAGGCTCTGAGATCTCAGTGCTCCAAAGTCCCTTGGC 1119
DB 1054 AAAAGATGAGCATGATTTGATCTCAAGAGAGGAGTGAAGCAATGGTGCAGACTATC 1113
QY 1120 CTCATCTGATACAAACATCTCGACATCCGCGCTCAGTATTGGACCTTGGCTAAT 1179
DB 1114 CTGCAGAGAGGAGCAACCGATGATGCTGCGCTCCGATGATGATCACTGCTAATGGA 1173
QY 1160 GAATATCTTCATGACAGACGATCCGATGATGATGATGATGATGATGATGATGATGAT 1239
DB 1174 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1233
QY 1240 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1299
DB 1234 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290
QY 1300 GACCTGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1359
DB 1291 GCGCTGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350
QY 1360 GCTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1419
DB 1351 CCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1410
QY 1420 AAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479
DB 1411 GGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
QY 1480 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1539
DB 1459 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1518
QY 1540 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
DB 1519 TACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1578
QY 1600 GACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1654
DB 1579 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1633

RESULT 3
LOCUS B1824830 576 bp mRNA linear EST 04-OCT-2001
DEFINITION 6030373881 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174937 5',
mRNA sequence.
ACCESSION B1824830
VERSION B1824830.1 GI:15936380
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 576)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1435 row: 1 column: 10
High quality sequence start: 6
High quality sequence stop: 565.
Location/Qualifiers

FEATURES
SOURCE
1. 576
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:5174937"
/lab_host="BD10B"
/clone_id="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-Sport; Site: 1; NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC library."

Query Match 23.7%; Score 414.6; DB 4; Length 576;
Best Local Similarity 96.4%; Pred. No. 3.8e-107;
Matches 456; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 91 GGGGCTTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 150
DB 104 GGGGCTTTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 163
QY 151 GTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 209
DB 164 GTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 223
QY 210 TCCCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 269
DB 224 TCCCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 283
QY 270 GGGAGAACCCATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCA 329
DB 284 GGGAGAACCCATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCA 343
QY 330 TCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 388
DB 344 TCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 403
QY 389 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 448
DB 404 TGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 463
QY 449 TCCCGAGAGGCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCA 508
DB 464 TCCCGAGAGGCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCA 522
QY 509 CCTATGAGAGGCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCA 561
DB 523 CCTATGAGAGGCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCA 575

RESULT 4
LOCUS AY414459 1680 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens CES2 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY414459
VERSION AY414459.1 GI:39770421
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1680)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeleo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.D., Smitsky, J.O.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302
 REFERENCE 2 (bases 1 to 1680)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Clevello, D.R., Lu, F., Murphy, B., Ferris, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J., Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 source
 1..1680
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..21680
 /gene="CBS2"
 /locus_tag="HCMS231"
 ORIGIN
 Query Match 23.3%; Score 406.6; DB 9; Length 1680;
 Best Local Similarity 56.5%; Pred. No. 1,1e-104;
 Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;
 QY 116 GGAAACCAAGCGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGTGC 175
 DB 101 GGACCAACACACAGGGGGCAGGTGTGGGAGCTTGTTCATGTGAAGGGCGCCATGTCCG 160
 QY 176 CTGTGAACGTGTTCCTCCGAGTCCCTTTGTGCTGCCCGGTGGAGTCCCTGCATTTA 235
 DB 161 GGGTCCAAACCTTCCCTGGGAATTCATTGGCAAGCACTCTAGGTCCGTGCATTTG 220
 QY 236 CGAAGCCGAGCGCTGATCGCCCTGGGATTAAGTGGCAAGCAAGTCTTACCTTAAT 295
 DB 221 CACCCCTGAGCCCTCCGTAATCTTGAAGTGTGTGAGGATGAAACCAATCCGAGCA 280
 QY 236 TGTGCTTCCAGAACTCAAGTGTGCTTTAGA---TCMACATGCTCAAGTGCATT 352
 DB 281 TGTGTCTACAGGACCTCAACCGCAGTGAAGTCAAGATTTCTTAGCCAGTTCAACATGACT 340
 QY 353 ACCGGAATTTGGAGTGTCAAGAACTGCTTACTTGAACATCTATGCGCTGCCAG 412
 DB 341 TCCCTCCGATCCATGTGTGAGACTGCTGTACTTCAAGCATTTACACGCCGCGCCATA 400
 QY 413 CGGATACAGGCTCAAGCTCCCGCTTGTGTGTGCTTCCAGAGAGTGCCTTCAAGACTG 472
 DB 401 GCATGAAGAGGCTTAACTGCGGTGATGTGTGATTCACGGTGTGCGCTTGTTTTG 460
 QY 473 GCTCAGCTCCATCTTTATATGAGGTCCGCTGTGCTGCTTATGAGAGCGTGTGTTGTG 532
 DB 461 GCATGCTTCTTGTATGATGTTCATGCTGTGCTGCTTGAAGAACGTGTGTGTG 520
 QY 533 TCGTCAGTACCGGCTAGGAATTTGGTCTTTCACCAACATGGAATCAGATCTCCG 592
 DB 521 TCATTCAGTACCGGCTAGGCTGTCTGTGGGCTTCTTCAAGCATGAGAACACCAACCG 580
 QY 593 GGAATCGGCGCTTGAAGAGCAAGGTGTGCTGTCTGTGGGTCCAGAAACATCGAGT 652
 DB 581 GCAACTGGGGCTTACTGGACCAAGTGTGCTGACCTAGGTGTCCAGCAAAATTCGCC 640
 QY 653 TCTTCGGTGGGAACCCAGCTGTGTACATCTTTGGGAGTCCGCGGAGCCATTAAGT 712
 DB 641 ACTTTGAGGCAACCTGACCGTGTCAACATTTTGTGCGAGTGTGCGGTGTGCGAGTG 700
 QY 713 TTTCTAGCTTAATCTGTCTCCATGAGCAAGGCTTAATTCACAAGCCATCATGAGA 772
 DB 701 TGTCTTGGCTTTGTGTGTCTCCCATATCCAGAACTCTTTCACGAGGCAATCATGAGA 760
 QY 773 GTGGGGTGGCCATCATCTTACTGTGAGGCGCAATGATTAAGAGAGTGAAGACCTGC 832
 DB 761 GTGGCGTGGCCCTCTGCTCCGCGCTCATGTGCGAGCTCAGCTGA-----TGTATCTCCA 814

QY 833 AGGTGTGACATTTCTGTGTGAACATGCGCTCAGACTGTGAGGCCCTGTGAGTCC 892
 DB 815 CGGTGTGGCCAACTGTCTGTGCTGTGACCAAGTTGACTGTGAGGCCCTGTGAGTCC 874
 QY 899 TGAAGCAAAACCTCCCAAGAGAGTGTGACCTTCAAGCAAAACAAAGTCTTCACTC 952
 DB 875 TCGGGGCAAGATTAAGAGAGATTTTGAATTTAACACCTTTCAAAGATGATCCCCG 934
 QY 953 GAGTGTGATGTGCTTCTTCTTCTTAATGAGCCCTGTGATCTTAATTTCTCAGAAAGCAT 1012
 DB 945 GAGTGTGATGTGAGGTCTTCTGCTCCAGGCAACCCAGAGACTGTGCTCTGCGACT 994
 QY 1013 TTAAGCAATTCCTTCCATCATGAGTCAATTAACAGAGTGTGCTTCTGTGCTTA 1072
 DB 995 TTACGCTGTCCCTAGCAATGTGTGTGTCACAAACATTAATGTGCTGTGCTATCCCA 1054
 QY 1073 TGA-----AGGAGGCTCCTGAGATCCCTCAGTGTCCCAACAGTCCCTTGGCC 1120
 DB 1055 AGGTGATGAGATCTATGATATCCAGAAAGAAATGACAGAGAGCCCTCCAGGCTGCTC 1114
 QY 1121 TCCATCTGATACAAACATCTCTGACATCCCGCTCAGTATTTGACCTTGTGCTAATG 1180
 DB 1115 TGCAAAATGTTAACGCTGTGATGTGTGCTCTTACATTTGTGACCTGTGAGGAGG 1174
 QY 1181 AATACCTTCATGACAGACATCTCCTGACTGAATCCGAGACAGTCTTGTGACTTGTG 1240
 DB 1175 AGTACATGTGGGACAAATGGGAGATCCCAAGACCTTCCAAAGGCAAGTTCAGAGATGATG 1234
 QY 1241 GAGATGTGCTTCTTGTGTGCTGCTGCTGACATGATCAAGCTGATATCAAGAGATGCTGTG 1300
 DB 1235 CGGATCAGATTTGTATCTCTGCACTCCAGTGAACATTTTCACTG---TTCCGCG 1291
 QY 1301 CACTGTCTACTTATGATGTTTGGCAACGCGCTCAAGTCTTTGAAGACAGAGCCG 1360
 DB 1292 CCCGTGTACTTCTACAGATTCAGATCAGCCAGCTGCTCAAGAACATCAGGCGAC 1351
 QY 1361 CTTTGTCAAGCGGACCAAGCTGATGAAGTCCGTTTGTGTGCTGTGCTTCTGCTGA 1420
 DB 1352 CGCAGATGAAGGACAGACATGTGTATGAGTCTTCTTGTTCAGAAATTTCTTTGGGG 1411
 QY 1421 AGGGGACATTTGTAATGTGGAAGAGCCAGAGAGAGAGATTAAGTCAAGCCGGAAGA 1480
 DB 1412 GCAACTACATTAATTT-----CACTAGGAAGAGAGAGCATGAGAGAGAGA 1459
 QY 1481 TGAATGAATCTGAGGCTACTTGTGCTGAAACCGGAAATCTTAATGGAACGACTGTCTC 1540
 DB 1460 TGAATGAATGACTGGGCAACTTTGCGAAGAAATGGGAACCCCAATGCGAGGCTGTCCAC 1519
 QY 1541 TGTGCGACGCTTAATCTGACTGAGCAGATCTCCAGCTGAGCTTGAACATGAGCTGTG 1600
 DB 1520 ACTGGCGCGTGTTCGACAGAGAGAGACATTAAGTCAAGCTGAACCTTACAGCTGCGGTG 1579
 QY 1601 GACAGAGACTCAAGAACCGCGGGGTGATTTTGAACGACCAACATCCCC 1651
 DB 1580 GCCGGGCTCTGAAGGCCCAAGGCTCAAGTCTGGAAGAAAGCGGTGCCCC 1630

RESULT 5
 AK033563
 LOCUS
 DEFINITION
 AK033563 2661 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030624L02 product:similar to CARBOXYPESTERASE PRECURSOR (EC 3.1.1.1) (AII-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE) [Mesocricetus auratus], full insert sequence.
 AK033563
 VERSION
 AK033563.1 GI:26329244
 KEYWORDS
 HTC, CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1

AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sun, L., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 12026611
PUBMED 12026611
REFERENCE 6 (bases 1 to 2661)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kozuka, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Ohno, M., Ohse, N.,
Okazaki, Y., Nishi, K., Nomura, K., Nunezaki, R., Ono, M., Ohsato, N.,
Okazaki, Y., Saito, K., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source 1. 2661
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:9030624L02"
/db_xref="taxon:10090"

/clone="9030624L02"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
25..1704
/note="unnamed protein product; putative
similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1)
(ALI-ESTERASE) (B-ESTERASE) (MONOBIOTINASE) (COCAINE
ESTERASE) (PROCAINE ESTERASE) (METHYLBIOTINASE)
[Mesocricetus auratus] (SPR|035533, evidence: PASTY,
83%ID, 100%length, match=1677)"
/codon_start=1
/protein_id="BAC28361.1"
/db_xref="GI:26329245"
/translation="MPLIKLGLNANVACGVLLVLYHVGODSASPIRNTHTGQVRGS
LVHVDIAHTFHTGIPAPCPVGPPLAPAPPEAPPEPSPGVRDGTSHNMCLQDNLM
GSEDLKMNLLPLPISMSDELTYNTIYVPAHAHESNLPMVWVHGGALVTGMASMD
GSMULATDDVVVAIIOYRLGVIFGFPSTDDHAKGMWGLIDVWALRWQONIVHGVAN
PDRVTIFESAGSIVSISVSPMSOGFHAIGMESGVAVLPDLSSSEMVHRIVAN
LSGCAVVSFTLMCLRGKNEAMLAIVKFKIIPGVYDGEFLPHQPOLMELPPECGLMER
VPSIIGINDRYGRIPLTPMDPAQKIBETITKTPAVYKSTALKNMLPPECGLMER
YMGDEDEFTLQAFREKMGDFMTVLPALQVAFHFORSHAPYTFEFQRPSPFDPR
PYVADHGDDEIPLVFGYQGNIKLPTREEBQLSRMKRYRANFARHGNPNSGULPYW
FVMDHDEQYLQIDLOPSVGRALKARRLOFWTKTLPQKIOELKGSQERHKEK"
2634..2639
/note="putative"
2661
/note="putative"
ORIGIN
polyA_signal
polyA_site
Query Match 22.8% Score 397.8; DB 3; Length 2661;
Best Local Similarity 56.1%; Pred. No. 4.2e-102;
Matches 872; Conservative 0; Mismatches 647; Indels 36; Gaps 5;
115 AGGAACACCAAGCTGGAGTGAATTCAGGCAAGCAAGTCACTGTGTGGAAAGCCCTGTG 174
124 AGGAACACACACAGGCAAGCTGGAGTGAAGGACGCTGTGCACGTGAAGACACTGACATC 183
175 CCTGTGAACGCTGTCTCTGAGATCCCTCTTGTCTGCTCCCGCTGGATCCTGTGGATT 234
184 GCTGTTCATACCTCTCTGGGAATCCCTCTTGCAGACCTCTGTGAGCACTGTGGCTTT 243
225 AGGAACCGGACCGCTGATTCGCGCTGGATTAATTGCGAAGAACCACTCTCACTCAT 294
244 GCACTCTTGAGGCGCTGAGGACATGAGTGGGTGAAGGATGGAATCTCACTCATCAAA 303
295 TTGTGCTTCAGAACTCAGAGTGGCTGCTCTTA--GATCAACATGCTCAAGGTGAT 351
304 ATGTGCTCAAAATGAACATTTGATGAGTTCAGAGATCGAAGATGAAATGAACTTGATC 363
352 TAACCGAAATTCGAGATGTCAGAGATGCTCTTACTTCACTGAACATTTATGCGCTCCAC 411
364 CTGCTCCCATCTCTATCTGAGAGATGCTGTATCTCAACATCTATGTATCAACCCAT 423
412 GCCGATACAGGCTCAACCTCCCGCTCTTGTGTGTTCCAGAGAGTGCCTTCAAGAT 471
424 GCCATGAGGGTCTTAACTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 483
472 GGCTCAGCTCCATCTTTGATGAGTCCCGCTGCTGCTGCTTGAAGAGATGCTGTGTG 531
484 GGCATGCTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
532 GTGCTCAGTACCGGCTGAGAAATTTGTTCTTCAACACATGAGGATCAGATGCTCG 591
544 GCTATCCAGTACCGGCTGAGTCTTCTGGAATTTTCAACACATGAGGATCAGATGCTCG 603
592 GGGACATGAGGCTTCAAGACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
604 GGCATGAGGATCTGAGACCAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
652 TTCTTGGTGGGAGCCCAAGCTGCTGATCACTTTGGCGAGTCCGCGGAGCCATAAGT 711

Db 664 CACTTGGAGCAACCTGACCGAGTCAACCATTTTGGAGAGTCAGCGGTGAGCAAGT 723
 Oy 712 GTTCTAGTCTTAACTGTCCTCCATGGCCAAAGGCTTATCCCAAAAGCATCATGAG 771
 Db 724 GTGCTTCAATGTTGTGTCTCCCATCTCCAGGAGACTCTTCCAGGTGCTCATGAG 783
 Oy 772 AGTGGGTGGCCATCATCTCTTACCTGGAGGCCATGATTTATGAGAAAGATGAGAC 831
 Db 784 AGTGAAGTACTGTGCTCCCTGATCTTATCTCCAGGCTCCCTGAG-----TGTTCAC 837
 Oy 832 CAGGTGTTGACAAATTTCTGTGTGTAAATGCGTCAAGCTCTTAAGGCGCTGTGAGTGC 891
 Db 838 AGAATGTGGCAATCTATCTGTGCTGTGACGTGTGAACTCAAGACCCCTGATGTCTGC 897
 Oy 892 CTGAGGACAAACCTCTCAAGAGAGCTGTGACCCCTCAGCGAGAAACAAAGTCTTCACT 951
 Db 898 CTAAAGGCGCAAGAAATGAAGAGAGATGTGGCTATTAAATGAGTCTTCAAAATCATCTCT 957
 Oy 952 CGAGTGTGATGTGTCTTCTTCTTAAAGAGCTCTAGATCTATTTGTCTAGAAAGCA 1011
 Db 958 GGTGTGTGATGAGAGATTCTCAACCAAGCATCTCAGAGAGTGTGCTCTTAAGAT 1017
 Oy 1012 TTTAAAGCAATTCCTTCCATCCGAGTCAATACCAAGATGTGGCTTCTGCTGCT 1071
 Db 1018 TTTCACTCTGTCTCCAGCATCATGTGTATCAAAATGAAATATGTGTGAATCTTCTCT 1077
 Oy 1072 ATGAAGAGAGCTCTCT-----GAGATCTCTAGTGTGCTCCCAACAGTCCCTTGCC 1119
 Db 1078 AGCATCATGAGCCAGCTCAGAAATATAGAGAAATACAGAAAGACCTGCTCACTGTT 1137
 Oy 1120 CTCACATGTATCAAAACATCTCTGACATCCCGCTCAAGTATTTGCACTTGTGCTAAT 1179
 Db 1138 CTAAAGAGCAGCCCTTAAATATGATGCTGCTCTGATGTGTGACCTGTCTAATGAG 1197
 Oy 1180 GAATATCTTCATGACAGCACTCCCTGACTGAATCCGAGACAGTCTTCTGAGCTTGCCT 1239
 Db 1198 GAGTCAATGGGGGAGACATGAGAGCCAGAGACTCTGCAAGACAGTTCAGAGAGATGAG 1257
 Oy 1240 GGAATGTCTTTTGTGTGCTCTGCACTGATCAGAGTCAATATCAGAGAGATGCTGT 1259
 Db 1258 GGGGACTTCATGTTGTGTATCTCTGCTCAGTCAAAATGAGACATTTTCA---ACGTTCCAT 1314
 Oy 1300 GCACCTGTCTACTTATGAGTTGGGACCGGCTCAGTGTCTTGAAGACAGAGCGG 1359
 Db 1315 GCTCTGTCTACTTATGAGTTCAACATGAGACCCAGCTTCTCAAGATTTTCAGGGCA 1374
 Oy 1360 GCTTTGTCAAGCCGACGAGCTGATGAGTCCGCTTGTGTGCTGAGTGCCTCTG 1419
 Db 1375 CCTATGTAAGGCTGACCATGATGATAATTTCTTGTCTTGTGCTACCACTTCTG 1434
 Oy 1420 AAGGGGAGCATTTGTATTTGGAAGAGCCAGAGAGAGAAATTAAGTCTGAGCCGAG 1479
 Db 1435 AACATTAATCTCTTA-----CACTGAGAAAGAGAGCAACTTAAGAGAGAG 1482
 Oy 1480 ATGATGAATTAAGTGGCTACCTTTGTGCTGAGACGGGAAATCTTAATGGAAGAGACTGTCT 1539
 Db 1483 ATGATGAAGTACCGGGCCAACTTTGACAGCGAGAGAGCCCAAGAGAGGCTGTACCC 1542
 Oy 1540 CTGTGGCAGATTATTAATGATGAGAGAGTCACTCAGCTGAGCTTGAACATGAGCTC 1599
 Db 1543 TACTGGCTGTGATGAGACATGATGAGAGTACCTGAGCTGAGCATTCAGCTTCTGTG 1602
 Oy 1600 GGAAGAGAGCTCAAGAAACCGCGGTGATTTTGAACAGACCATCTCCCTGTA 1654
 Db 1603 GGTGAGCCCTGAAGGCGCAAGAGGCTCAGTTCTGAGCAAGACTCTGCCCCAGA 1657

RESULT 6
 AK077248 2687 bp mRNA linear HTC 03-APR-2004
 LOCUS
 DEFINITION Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
 RIKEN full-length enriched library, clone:5031415B19
 Product:similar to CARBOXYL-ESTERASE PRECURSOR (EC 3.1.1.1)

COMMENT

(ALL-ESTERASE) (B-ESTERASE) (MONOBUITYRASE) (COCAINE ESTERASE)
 (PROCAINE ESTERASE) (METHYLBUTYRASE) [Mesocricetus auratus], full
 insert sequence.
 AK077248
 AK077248.1 GI:26346111
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, U.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6
 (bases 1 to 2687)
 Adachi, U., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, U., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saitoh, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Yurimatsu, M. and Hayashizaki, Y.
 Direct SubMISSION
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genome Sciences Center (GSC),
 Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>

FEATURES

source

1.2687
/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:5031415B19"

/db_xref="taxon:10090"

/clone="5031415B19"

/sex="female"

/tissue_type="ovary and uterus"

/clone_1ib="Riken full-length enriched mouse cDNA library"

/dev_stage="11 days pregnant adult"

53..1732

/note="unnamed protein product; putative
similar to CARBOXYL ESTERASE PRECURSOR (EC 3.1.1.1)
(ALI-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE
ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
(Mesocricetus auratus) (SPTB|O35533, evidence: PASTY,
83%ID, 100%length, match=1677)"

/codon_start=1

/protein_id="BAC36707.1"

/db_xref="GI:26346112"

/translation="MPLVYLGLMNAVACGVLLVLAHVGDSASPIRNTHTGVRGS
LVHVDTDIAVHTPLGIPPAKPPVGRPLRPAEPBPSGVDGSHRMCLQNDLM
GSEDLKMMNLILPISMSDELILNIYVPAHAGSNLPEVWVHIGALITVMSMYD
GSMILAKMDELIVVVAIOYRLGVLPFSTGPAHAGNGLVDVLAHVMQNDVHSGM
PDRVITFESAGTSVSHVAPMSQFLGPAHAGNGLVDVLAHVMQNDVHSGM
LSGCAVNSSETMCLRGKNBAEMLAINKFPIIPGVNDEFLPKHPOBMSXDFPH
VPSILGINDEYGMILPTIMDAOKIETRTPLPAVLKSTLAKMLPBCDILME
YMGEDBEDPDLQAQFREKMGDFMFIPLAQVHFORSHAPVPIYERFORHPSFQDPR
PVKADHDEILPLVFGYQGNLKLPTBEEQDSRIKMTANPAHAGNPSGGLPYW
PVMHDHQLQDIDPVSGRALKARLQFWITLPLQKIQELNGSOERHKL"

2664..2669

/note="putative"

2687

/note="putative"

polyA_signal
polyA_site

/note="putative"

ORIGIN

Query Match 22.7%; Score 396.2; DB 3; Length 2687;
Best Local Similarity 56.0%; Pred. No. 1.2e-101;
Matches 871; Conservative 0; Mismatches 648; Indels 36; Gaps 5;

QY 115 AGGAACACCAAGCTGGATGATTCAGGGCAAGCAAGTCACTGTCTGGAAACCCCTGTG 174
DB 152 AGGAACACACACAGGCGAGGAGAGGACAGCTTGTCCACGTGAAGACACTGACATC 211
QY 175 CCTGTGAAGCTTCTCGGAGTCCCTTTTCTGCTCTCCCGCTGGGATCCCTGGCATTT 234
DB 212 GCTGTTCATACCTTCTCGGAAATCCCTTTGCAAGCCCTCTGTAGAACACTGCGCTTT 271
QY 235 ACGAACCCGACGCTGCATCGCCCTGGGATTAATTGCGAGAAAGCACTCTTAACCTAAT 294
DB 272 GCACTCTCTGAGGCGCTGAGCCATGAGTGGGGTAGGAGTAGGACCTTCACTCAAC 331
QY 295 TTGTGCTTCAGAACTCAAGTGTGCTCTTA--GATCAACACATGTCTCAAGGTGAT 351
DB 332 ATGTGCTTACAAATGACATTTGATGGGTCAAGGATCTGAAGATGATGAACCTTGATC 391
QY 352 TACCCGAATTTGGAGTGTCAAGAACTGCTCTCACTGAACATCTATAGCGCTGCGCC 411
DB 392 CTGCTCCCATCTCTATGTCTGAGACCTGCTATCTCAACATCTATATACCAAGCCAT 451
QY 412 GCGATACAGGCTCCAAAGCTCCCGCTGTGTGTGGTCCAGAGAGTCCCTTCAAGACT 471
DB 452 GCGCATAGAGGCTTCTTAACTCTGTGATGTATGATCATGTGTGTGATCACTCACTGTT 511
QY 472 GCGTCAAGCTCATCTTTGATGGTCCGCTGCTGCTTATGAGGACGTGTGCTGTGTG 531
DB 512 GGCATAGGCTTCAATGATGATGATCCATGTCTGGACGAACCTGAGGATGTGGTGGTGTG 571

QY 532 GTGTCCAGTACCGGCTAGGAATATTGTGTTCTTACACACATGGGATGAGATGCTCCG 591
DB 572 GCTATCCAGTACCGGCTAGGATTTCTGGATTTCTTACAGACTGAGAACAGATGACAA 631
QY 592 GGAACCTGGGCTTCAAGACCAAGTGTGCTGTCTGTCTGGGTCCAGAAACATGAG 651
DB 632 GGAACCTGGGATATCTGGACCAAGTGTGCTGTCTGTCTGGGTCCAGAAACATGAG 691
QY 652 TTCTTGGTGGGATCCCAAGCTGTGACCACTTTTGGGAAATCCGCGGACCATTAAT 711
DB 692 CACTTGGAGGCAACCTGACCGAATCACTTTTGGAGATCAAGCAAGTGCACAAAT 751
QY 712 GTTTCACTTATATCTGTCTCCCAAGGCTTATTCACAAAGCCATCAATGAG 771
DB 752 GTGTCTTCACTTGTGTGTGTCTCCCAAGTCTCAAGGATCTTCAAGTGTGATGAG 811
QY 772 AGTGGGTGGCCATCATCTTACCTGAGGCGCATGATATGAGAAAGTGAAGACTG 831
DB 812 AGTGAAGT 865
QY 832 CAGTGTGTGACATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 891
DB 866 AGAATAGTGGCCAACTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 925
QY 892 CTGAGACAAACCTCCCAAGAGCTGCTGACCTCAGCCAGAAACAAAGTCTTCACT 951
DB 926 CTAAAGGCAAGAAATGAGACAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 985
QY 952 CGAGT 1011
DB 986 GGT 1045
QY 1012 TTTAAAGCAATCTTCCATCATCTGGAATCAATCAAGATGTGTGTGTGTGTGTGTGT 1071
DB 1046 TTTCACTGT 1105
QY 1072 ATGAAGAGGCTCTCT-----GAGATCTCAAGTGTGTGTGTGTGTGTGTGTGTGT 1119
DB 1106 ACGATCATGACCCAGCTCAAGAAATAGAGAAATACAGAAAGACCTGCACTGTGT 1165
QY 1120 CTTCATCTGTATCAAAACATCTGTGACATCTCCGCTCAAGTATTTGCACTTGTGTGT 1179
DB 1166 CTAAAGAGACAGCCCTTAATAATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1225
QY 1180 GAATCTTCTATGACAGACATCTCTGACTGAATCCAGACAGTCTTGTGACTTGTCTT 1239
DB 1226 GAGTACATGAGGAGACATGAGGACCCAGACTCTGCAAGCAAGTTCAGAGATGAG 1285
QY 1240 GGAATGTGTCTTTGT 1299
DB 1286 GGGAGCTTCAATGT 1342
QY 1300 GCACTGTCTACTTCTATGATTTGCGACCGGCTCAAGTCTTGTGAAGACAGAGCG 1359
DB 1343 GCTCTGTCTACTTCTATGATTTGCGACCGGCTCAAGTCTTGTGAAGATTTGCGGCA 1402
QY 1360 GCTTTGTCAAAAGCGACACGCTGATGAAGTCCCTTTGTGTGTGTGTGTGTGTGTGT 1419
DB 1403 CCTATGTGAAGGCTGACCATGTATGAATTTCTTGTGTGTGTGTGTGTGTGTGTGT 1462
QY 1420 AAGGGGACATTTGATATGTGAAGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1479
DB 1463 AACATTAATCTCTTA-----CATGTGAAGAGGAGGAGGAGGAGGAGGAGGAGG 1510
QY 1480 ATGATGAATAATCTGGGCTACCTTGTCTGAAACGGGAAATCTTAATGAGAACAGCTGTCT 1539
DB 1511 AATAATGATGATCTGGGCACTTTGCAAGGACGGGAAACCCAAACAGTGAAGGTCTAACCC 1570
QY 1540 CTGTGGCAGCTTAAATCTGATGAGAGTACTTCAAGCTGTGACTTGAACATGAGCCTC 1599
DB 1571 TACTGGCTGTGATGAGACATGATGAGAGTACTTCAAGCTGTGAGATCATCAGCTTCTGTG 1630

QY 1600 GGACAGAGACTCAAGAACCGGGGTGATTTTGGACCGACCATCCCCCTTA 1654
Db 1631 GGTGAGCCCTGGAAGGCCAGAGCTGCTGACGACGACGATCTGCCCCAGA 1685

RESULT 7
AK040349

LOCUS
DEFINITION
AK040349 4927 bp mRNA linear HTC 03-Apr-2004
Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched
library, clone:A430088E12 product:similar to CARBOXYESTERASE
PRECURSOR (EC 3.1.1.1) (ALT-ESTERASE) (B-ESTERASE) (MONOBUTYRASE)
(COCAINE ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
(Mesocricetus auratus), full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK040349
AK040349.1 GI:26087790
HTC; CAP trapper.
Mus musculus (house mouse)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M.,
Sun, N., Ishi, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
RIKEN integrated sequence analysis (RISA) system--384-format
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
20530913
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
11076861
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
11076861
6
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirose, T., Hirose, T.,
Hori, F., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Murata, M.,
Nakamura, Y., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnishi, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shikawa, A., Shikata, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takahashi, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Yamamoto, M., and Hayashizaki, Y.

TITLE
JOURNAL
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

COMMENT
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>.
URL: <http://fantom.gsc.riken.jp/>.
Location/Qualifiers
1. 4927
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A430088E12"
/db_xref="taxon:10090"
/clone="A430088E12"
/issue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_abbrev="0 day neonate"
25. 1711
/note="putative
similar to CARBOXYESTERASE PRECURSOR (EC 3.1.1.1)
(ALT-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE
ESTERASE) (PROCAINE ESTERASE) (METHYLBUTYRASE)
(Mesocricetus auratus) (SPR103553, evidence: FASTY,
70.7%ID, 99.2%length, match=1672)"

misc_feature
25. 1711
/note="putative"
polya_signal
4912..4917
/note="putative"
polya_site
4927
/note="putative"
ORIGIN
Query Match 22.2% Score 387.6; DB 3; Length 4927;
Best Local Similarity 56.1%; Pred. No. 4.3e-99;
Matches 890; Conservative 0; Mismatches 659; Indels 37; Gaps 7;
QY 90 TGGGCTTCTGCTGAAGGGCCACAGAGAACACACAGGCTGGATTCAGGCAAGCA 149
Db 105 TGAATCATGAGAGGCCAGGCCCATGAGAAACACATCTGACAGCTCGAGGCAAGTT 164
QY 150 AGTCACGTGCTGGAAGCCCTGCTGTAACGTGCTGGAAGTCCCTTGGCTGC 209
Db 165 TGTCACTTGAACAGACATGAACCTGAGCCCATTAATCTCTGGGATCCCTTGC 224
QY 210 TCCCGCTGGAGATCCCTGATTTAGAAC-CCGAGGCTGCATGCGCTGGATTA 268
Db 225 GCTTCTGTAGAGACATGCTGCTTGCACCTTCTGAGGCCCTGAGCATGAGTGT 284
QY 269 TCGAAGAGCCATCTTCAACCTTAATTTGCTCGAGAACTCAGAGTGGCTCTT 328
Db 285 TGAAGATGGAAGCTCAGACGCGGCGCATGTGTCAAAATGATATTAAGATTAG 344
QY 329 A---TCAACATGCTCAAGGTGCATTAACCGAATTGGAGTGCAGAGCTGCT 385
Db 345 AGGCTTGAAGATTAATAAAGATGATCTGCTCCCTTCTATGTCTGAGACTGCT 404
QY 386 ACTGAACATCTATGAGCTGCTGCGCAAGCGATACAGGCTCCAAAGTCCCGCT 445
Db 405 ATCTCAACATCTACACACAGCCAGCCGCGAGAGGCTTAACTTCTGATGATGT 464
QY 446 GGTTCAGAGAGGTGCTTCAAGACTGCTCAGCTTCATCTTTGATGAGGTCCG 505
Db 465 GGCCTCATGTGTGTGAGCTGTTGACGAGCATGAGCTTCATGATGATGATCCG 524
QY 506 CTGCTATGAGAGCGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565
Db 525 CAGCCATGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 584

566 TCACACATGGGATGATGATGCTCCGGGAACTGGGCTTCAAGGACAGGTGCTGCTC 625
 585 AAGACATCTGGAGATGAGAGAGGCGACAGGCGAATGGAGATTCTTAAGCAAAACAGCTGCC 644
 626 TGTCTGGGTCAGAGAAACATCGAGTTCTTGGTGGGAGCCCGAGCTCTGTACATCT 685
 645 TACACTGGGTTCAGAGAAACATTTGGCACTTTGGAGGCAACCTGATCTGTGACATTT 704
 686 TTGGCGAGTCCCGGAGACCATTAATGTTTCTAGCTTATATCTGTCTCTCCATGGCAAG 745
 705 TTGGCGAGTCCCGGAGACCATTAATGTTTCTAGCTTATATCTGTCTCTCCATGGCAAG 764
 746 GCTATTTCACAAAGGACCATCATGAGAGTGGGGTGGCCATATCCCTTACCGGAGGCC 805
 765 GACTCTTCATAGAGCCATCATGAGAGTGGGGTGGCCCTGCTGCTTACATTAATCTG 824
 806 ATGATTATGAGAGAGTGAAGACCTGACAGTGTGCAATTCTGTGTGAACAATGCGT 865
 825 ACTCCCTGAGATGA-----TCTTCAAAAGTGGCAAACTTATCTGTTGTGAACCT 878
 866 CAGACTGTAGAGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGCTGCTGACC 925
 879 CAATCTAGAGAGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGAGCTGCTGACC 938
 926 TCAGCCGAAAGAAACAAAGCTTCACTGAGTGTGATGAGTCTTCTTCTTATAGAC 985
 939 TGACCAAGGCTTCAAGGTTTATGCTGCTGTGTGTGAGGAAATTTCTTACCAACATC 998
 986 CTCTAGATCTATGTTCTCTGAGAAAGCATTTAAAGCATTTCTTCCATCATCGAGTCAATA 1045
 999 CCAAGAGTCTGCTGCTCTGTCTGATCTTCACTTCACTTCCAGATTTTGTGTCAACA 1058
 1046 ACCAGAGTGTGCTCTGCTGCTCTGATGAAG-----AGCTCTCTGATGCC 1093
 1059 ATGATGATGATGCTGATGATATCCCTAAGATCTTCAAAATTTCTCAGACAATTAAGAA 1118
 1094 TCAATGCTGCAACAAGTCCCTGCGCCCATCTGATCAAAACAATCTGACATCCCGC 1153
 1119 TAAATAGAAATTAATCTGAAGCTATTAAGAGATCAACAAGCAAAATGATGCTGCTT 1178
 1154 CTCAGATTTGACCTTGTGCTGCTGATGAATATCTTCCATGACAGACATCCCTGACTGAA 1213
 1179 CTGAGTGTGAGACTGCTGCTAATTAAGAGATCTTAAGAGACATGAGGACCCCTGGAAC 1238
 1214 TCCGAGACAGTCTTCTGAGCTTGTGAGAGTGTCTTGTGTGCTCTGATGATCA 1273
 1239 TCCAAATGCACTTCAAGAGATGATGAGGGAATCTTAAATTAATATCCCTGATCCAA 1298
 1274 CAGCTGATATCAAGAGATGCTGTGATGACCTGCTATCTTCAATGATTTGGGACCGGC 1333
 1299 TAGACGTTTTCAAGG--TTCCATGCTCTGCTGCTTCTTAAGATTCAGCAATAGGT 1355
 1334 CTCAGTCTTGAAGACAGAGCGGCTTTGTCAAGCGGACCAAGCTGATGAATGCC 1393
 1356 CCAATTAATCAAGATTTCAAGGCAATGAGCATTTAAGGCTGACCAAGCTGATGAGCTT 1415
 1394 GCTTTGTGCTGCTGCTTCTTCTGAAAGGAGCAATTTGTTTGAAGAGGACGACGG 1453
 1416 ACCTTAATCTT---TGATCTCTTCTTCTGAGGCTCTCAAAATTAAGCTT-----CACTG 1463
 1454 AGAGAGAGAAATTAATGAGCCGGAAGATGATGAATATCTGGGCTTACCTTTGCTGCAACG 1513
 1464 CGAGAGAGAAATCTGAGAGAGAAATGATGAATATCTGGGCTTACCTTTGCAATGATG 1523
 1514 GGAATCTTAATGAGAGAGAGCTGCTCTGTGGCCAGCTTATTAATGATGATGAGTACC 1573
 1524 GGAATCTTAATGAGAGAGAGCTGCTCTGTGGCCAGCTTATTAATGATGATGAGTACC 1583
 1574 TCCAGCTGAGCTTGAACATGAGCTCTGAGACAGAGCTCAAAAGACCGGAGTGAATTTT 1633
 1584 TGACAGCTGAGATCAACCTTCTGTGGGCTCAATGCTTGAAGGACGAGAGGCTGCAATCT 1643

QY 1634 GGACGAGCACCATCCCTGATCTG 1659
 DB 1644 GGACCAAGATCTTCCCAAGAGATG 1669
 RESULT 8
 AK078953
 LOCUS
 DEFINITION
 AK078953 1958 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male cecum cDNA, RIKEN full-length enriched
 library, clone:9130231C15 product:similar to LIVER CARBOXYLSTERASE
 PRECURSOR (EC 3.1.1.1) [Mesocricetus auratus], full insert
 sequence.
 ACCESSION
 AK078953
 VERSION
 AK078953.1 GI:26347654
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2
 99279253
 10349636
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subcloning of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 3
 20499374
 11042159
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Komuro, H., Akiyama, J., Nishi, K., Kikunaka, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 4
 20530913
 11076861
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 405, 685-690 (2001)
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1958)
 Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanakawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Komodo, S., Komuro, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakezume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takehashi, F., Takeuchi-Akita, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome

|||||
Db 1524 ATGAAGTACTGGCCCAATTTCGACAGAAATGGGAACCCCAACGTGAGGCTGTACCTCC 1583
Qy 1543 TGGCCAGCTTAAATCTACTGACGAGTACCTTCAGCTGAGCTTGAACATGAGCCTCGGA 1602
Db 1584 TGGCTGTGTGTGGACCATGATGACGTACCTGACGTGAGACCCAGCCTCTGTGGGT 1643
Qy 1603 CAGAGACTCAAGAACCCGGGGTGGATTTTGGACGAGCAGCCATCCCTCTGA 1654
Db 1644 CGAGCCCTGAAAGCCCAAGAGGCTGCACTTTCGACCAAGACTCTGCCCCAGA 1695

RESULT 9
LOCUS AK078879 1944 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030604P03 product:carboxylesterase 3, full insert sequence.
ACCESSION AK078879
VERSION AK078879.1 GI:26347580
KEYWORDS HTC, CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtration of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Suni, N., Ishi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillarary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
AUTHORS 5
The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
AUTHORS 6 (bases 1 to 1944)
Aduchi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Haseizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishi, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, Y., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.jp, URL:htp://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:htp://genome.gsc.riken.jp/
URL:htp://fantom.gsc.riken.jp/.
URL:htp://location/Qualifiers

FEATURES
source 1..1944
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9030604P03"
/db_xref="taxon:10090"
/clone="9030604P03"
/sex="male"
/cissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
64..1761
/note="unnamed protein product; carboxylesterase 3 (MGD|MGF:2148202, GI|NM_053200, evidence: BLASTN, 99%, match=1939)
putative"
/codon_start=1
/protein_id="BAC37439.1"
/db_xref="GI:26347581"
/translation="MGLPLILWLSLAACITAWKIPSPPVNTVKGKVLGKYNLSEFT
QVAVFLVGPPEKPLSLRFAFPQPAEWSVVKNTSTPPMCSQDAVGQVLSLFT
NRKNTPLPFSEDCYLNIYTPADLPALRSLRWVWVHIGGGLVVGASTYDGLALSAH
ENVVVTVIOYRLGIVGFSTGDEHSGWGHLDQVALLWVDNLNFGNGSGSVTF
GSSAGGSVSVTLVPLAKNLFPHRAISGSGVLTALITDVKPIAGLVATSGCCTT
TSAAVHNCRCOTREDELLETSLKILPLDILGNKESVPLPIYDGVLPKAPBEI
LAEKSFSTPIYVINKQEFWITITTLNGYPLAEGLDOKTANSLMKSYPFKIKSEN
MIPVAEKRIAGTDLTLKKDLFDQMLADVGVSVIVSRHDAQASTVWEYIR
ESFVAMPKAVIGDHGDEISFVFGSPFKDQASBEETLSQVWKFPMFANFANGNP
GGGLPHMPEYDQKEGYLKIKAISTQAORLKDKEVSFMELRANESAPSRHEVEL"
1922..1927
/note="putative"
1944
/note="putative"
polya_signal
polya_site
ORIGIN
Query Match 20.3%; Score 355.2; DB 3; Length 1944;
Best Local Similarity 55.1%; Pred. No. 6.3e-90;
Matches 870; Conservative 0; Mismatches 633; Indels 75; Gaps 6;
Qy 117 GAACACGAGCTGGATGATTCAAGGCAAGCAAGTCACTGCTGGGAAGCCCTGTGCC 176
Db 141 GAACACTGTTAAAGCAAGTCTCTGGGAAGTATGTCATTGGAAGATTCAACAGCC 200
Qy 177 TGTGAACGTGTTCTCTGGAGTCCCTTGTCTGCTCCCGCGCTGGGATCCCTCGATTAC 236
Db 201 TGTGCTGTGTTCTCTGGAGTCCCTTGTCCAAAGCCCTCTGTGCTCTTGAAGATTGC 260
Qy 237 GAACCGGAGCTGTGATGCGCTGGATTAATTGGAGAGACACCTCTACCTTAATT 296
Db 261 TCCACCAACGCTTGAGAGCCCTTGAGAGCTTCTGTGAAGAACACACCTCTACCGGCTAT 320
Qy 297 GTGCTTCAGAACTC-----AGATGCGTGTCTTTAGTCAACATGCTCAAGGTGCA 350
|||||

D	b	331	GTGCTCTGAGATGCTGTGTTGAGGGCAGGTGCTCTCAGAGCTTTTCAACCAAGAAAGGA	386
Q	y	351	TTACCCGAAATTCGGAGTGTCAAGAACTGCCTTCACTGAAACAATTATGCGCCCTGCGCA	410
D	b	381	GAACAATTCCTTTACAGTTTCTGAAAGACGTGCTCTCACTGAAATTTTAACTCTCTGCTGA	446
Q	y	411	CGCGGATACAGGCTCCAAAGCTCCCCGTGTGATGTGGTTCCAGAGAGTGCCTTCAAGAC	470
D	b	441	CTTGACAAAGAAACAGACAGACTAACAGTGAATGATGATGATTCATGAAAGTGAATGATGAT	500
Q	y	471	TGCTCAAGCTTCCATCTTTGATGAGGTCCGCCCTGAGCTCATAGGACGTGCTGTTGT	530
D	b	501	GGGGGAGCAATCCACTATGATGAGCTGGCCCTCTGTGCCATGAAATGTGGTGTGT	566
Q	y	531	GGTGTCCAGATACGGGCTAGGAAATTTGTTTCTTTCACCAATGGATACAGATGCTCC	590
D	b	561	GACCATTTCAAGTATGCTCTGGCACTGAGGAAATTTTTCAGACATGGGAGATTAACAAGTCC	620
Q	y	591	GGGGAATGGGCTTTCAGAGACCAAGTGGGCTCTGTCTCGGGTCCAGAAAGAACATCGA	650
D	b	621	GGGAAATGGGGGTCACTTTGGACAGATGATGCTGACCTAACGTGGTCCAGAACAACTTGC	680
Q	y	651	GTTCTTGGGTGGGGAACCCCAAGCTCTGTGACCATTTTGTGCGAGTCCGGGAGCCATAAG	710
D	b	681	CAACTTTGGGGGCAACCCAGGCTCGGAGCAATCTTTGGAGATGTACAGAGAGTTTCAG	740
Q	y	711	TGTTTCAAGCTTATATCTGTCTCCATGGGCAAGGCTTATTCACAAGAGCCATCATAGA	770
D	b	741	TGCTCTGTTCTTGTCTTGTCTCTCTTTGGCAAGAACCTTTTCACAGGCGCAATTTCTGA	800
Q	y	771	GATGGGGGTGGCCATCATCCCTTACCTGGAGGCCCATGATTTAGAGAAAGTAGGAGACT	830
D	b	801	GAATGTGTGTCCCTCACTGCTCTCTGTATTAACAAGATGTAAACCCATGTCTGTCT	860
Q	y	831	GCAGTGTGTGACATTTCTGTGTGTAAACAATGCTGAGACTCTGAGGCCCTGTAGAGTG	890
D	b	861	GGTTGTCTACTTTTCTGGGTGTAAACATACACATCAGCTGTATGTTCATTGTGCTGG	920
Q	y	891	CCTGAGGACA-----AAACCTCAAGAGAGCTG	918
D	b	921	CCAGAAAGACAGAGATGATCTACTGAGACCTCATAAAATTGAATCTTTTAACTGGA	980
Q	y	919	CTGAGCCCTGAGCGAGAAAACAAG-----TCTTTCATCTCAGTGTGTGATGTGTCTT	971
D	b	981	CTTACTTGGAAATCCAAAGAGAGCTATCCCTTCTCTCTACTGTGATGTGAAGAGTGT	1040
Q	y	972	CTTTCCTTAATGAGCCTCTAGATCTATTTGTCTCAGAAAGCATTTAAAGCAATCTCTTCAT	1031
D	b	1041	TCGCGCAAAAGCACCAAGAGAGATCTGGCTGAGAAAGATTTCAACAATGAGGCTATCACT	1100
Q	y	1032	CATCGGAGTCATTAACAAGAGTGTGGCTTCTGTCTGCTGCTTATGAAGAGG-----C	1082
D	b	1101	AGTGGCATCAACAGCAAGAGATTTGGCTGATCATTTCAAGCTTATGAGGCTATCACT	1160
Q	y	1083	TCCTGAGATCTCAAGTGTCTTCAACAAGTCCCTTGGCCCTCCATCTGATACAAAT---AT	1139
D	b	1161	CGCTGAAGCAAACTGAGACCAAGAGACGCTCAATCTCTTGTGTGAAGTCCATCCAAC	1220
Q	y	1140	CCTGACATCCCGGCTCAAGTATTTGACCTTGTGCTATGATATCTTCAATGACAGCA	1199
D	b	1221	ACTTAAATCTCTGAAATATGATTTCCAGTGTCTGCTGAGAAATTTAGAGGGAGCAGA	1280
Q	y	1200	CTCCCTGACTGAATTCGAGACAGTCTTCTGGAATTGCTTGGAGATGTGTTCTTTGTGT	1259
D	b	1281	TGACTTCAACAAAAGAAAGACTGTTCAGAGACTTGAATGCTATGTGTGTGTGTGT	1340
Q	y	1260	CCCTGCACTGATCAGGCTCGATATCACAGAGATGCTGTGCACTGTCTTCTATGA	1319
D	b	1341	CCCACTCAATGATTTGTGTCTCGAAGTCA CAGAGATGCTGAAGCTTCACTATATGATGA	1400
Q	y	1320	GTTTTGGGCAACGGGCTCAGTGTCTTGAAGAACAGAAAGCCGGCTTTGTCAAGCCGACCA	1379
D	b	1401	GTTTGATGTATGCCCAAGCTTTGTATGCGCCATAGAACCAAGGCAATTAATGAGAGCA	1460

QY	1380	CCCTGATGAAAGCCGCTTTGGTTGCGGGTGCGCTTCTCTGAAAGGGGGACATTTATATT	1459
Db	1461	TGGTATATGATCTTCTCAATATTTGGATCTCCATTTTTAA-----	1502
QY	1440	CGAAGAGCACGAGGAGAGAAATTACTAGACCGGAGATGATGAAATACTGGCTAC	1499
Db	1503	AGATGGTGCCTCAGAAAGAGAGACCAACTCAGCAAGATGGTATGAAATTTGGGCCA	1582
QY	1500	CTTGGCTGAAACCGGGAATCTTAAATGGGAAGACCTGTCTGTGGCCAGCTTAAATCT	1559
Db	1563	CTTTCCTGGAATGGGAACCCCAATGGTGGAGGGCTGCCCACTGGGCAGAAATGTGACA	1622
QY	1560	GACTGAGCAGTACCTCCAGCTGGACTTGAACATGAGCCTCGGACAGAGACTCAAGAAC	1619
Db	1623	GAAAGAGGATATCTGAAAGATTGGTTGGCTCAACTCAGGACCCAGAGCTGAAAGACAA	1682
QY	1620	GCGGTGATTTTTTGAC	1637
Db	1683	AGAAATGATTTTTGGGC	1700
RESULT 10	HSMB03013		
LOCUS	HSMB03013	2888 bp	linear
DEFINITION	Homo sapiens mRNA, CDNA DKFPZ434N0935 (from clone DKFPZ434N0935).		
ACCESSION	AL713761		
VERSION	AL713761.1		
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (Bases 1 to 2888)		
AUTHORS	Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R., Mewes, H. W., Weill, B., Aml, C., Oeinger, A., Fobo, G., Han, M. and Wiemann, S.		
CONSRPT	The German CDNA Consortium		
TITLE	Direct Submission		
JOURNAL	Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY		
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ). Email s.wiemann@kfz-heidelberg.de; Heidelberg/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFPZ434N0935) is available at Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneId=DKFPZ434N0935 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.		
FEATURES	location/Qualifiers		
source	1..2888		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="RZPD:DKFPZ434N0935"		
	/db_xref="taxon:9606"		
	/clone="DKFPZ434N0935"		
	/tissue_type="testis"		
	/clone_1ib="434 (synonym: hres3). Vector pSport1; host DH10b; sites NotI + SalI"		
	/dev_stage="adult"		
	/note="cardoxylesterase 2 isoform 2"		
	1..2888		
	/gene="DKFPZ434N0935"		
	954..2777		
	/gene="DKFPZ434N0935"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAD28531.1"		
	/db_xref="GI:19584507"		
	/db_xref="UniProt/Swiss-Prot:O00748"		
	/translation="MTATGSRPPTPTGSGPSPPTPTDPCVTPPTPTCTATTTCTCTTCTG		

```
/db_xref="UniProt/Swiss-Prot:000748"
/translation="MTAQSRSPTPTFPGPSQRTPLTPCPVQTPRLGKALIHCVTDPG
```

OPLGQQRVRRORETSEPTWMLHRLARLASAVACGLLLLVRCGQDSASPIRTHT
 GOVLGSLVHVKANAMGVOTFLGIIPKAPKPLGLRPARPAPPEPMSVWRGTHPAMLC
 ODLTBSEPLSOPMPTPSPDSMSDCLYLSTYTPAHSPREPSLPMVHIGGALVRC
 MASLYDSGMLALBNVVVVIYIRLGLVLPFESTGKHATNGWYLDQVAAHVRQONT
 AHFNGNPRVITFGESAGTSVSLVSTISQGLPHGALMESGVALLPELILASADVI
 STVVNMLASCDQVDEBALVGLRGSKEIILANKPFXKIPEVVGVFLPRHQBELA
 SADFPVPSIVGVNNNEFGMLIPKVMRIYDQKEMDREASQALQKMLFLALMLPTFG
 DLREBYIGDNDDPOTLOAQPOEMMAHMFVILPALOVAHFOCSRAPVYVEFOHOPSM
 LKNIRPPHMKADHVKPTREEROLSRKMVMYANPARNBNBEGILPHMFLPQOEBYL
 QNLQOPAVGRRLKARLPQPKALPQKIOLEBEPBRHTL

Query Match 19.8%; Score 345.8; DB 3; Length 2888;
 Best Local Similarity 54.8%; Pred. No. 3.6e-87;
 Matches 850; Conservative 0; Mismatches 617; Indels 84; Gaps 5;

116 GGAAACACAGGCTGGAGTGAATTCAGGGGACAGCACTGCTGCTGGGAAGCCCTGTGC 175
 1246 GGACACACACACAGGGGGCAGGTGTGGGGAGTCTTGTCCATGTGAAGGGCGCCAAATGCCG 1305
 176 CTGTGAACGTCTTCTCGGAGTCCCTTGTGCTGCTCCCGCTGGAGATCCCTGGCATTTA 235
 1306 GGGTCCAAACCTTCTCGGAAATTCATTGGCAAGCACCTCTAGGTCCGCTGCGCATTTTG 1365
 236 CGAACCCGAGGCTGCATGCCCTGGGATTAACCTTGGAGAGACCACTCTACCCCTAATT 295
 1366 CACCCCTGAGCCCTCGAATTTTGAAGTGTGTAGAGGATGGAACCACTCATCCGGCCA 1425
 296 TGTGCTTCAAGACTCAGAGTGTGCTCTTAGA---TCAGACATGTCTCAAGAGTGCATT 352
 1426 TGTGTCTACAGGACCTCAGCGAGTGAAGTCAAGTTCCTTAGCCAGTTTCACATGACCT 1485
 353 ACAGCAATTCGGAAGTGTCAAGAACTGCTCTTCTTCACTGAACCTATATGCGCTGCCCAAG 412
 1486 TCCCTTCCGACTCCTATGTCTGAGAACTGCTCTGTAAGCTCAAGCAATCAAGCCGGCCCAAT 1545
 413 CCGATPACAGGCTCAGAGCTCCCGCTGTGTGTGTGTCCAGAGAGTGCCTTCAAGACTG 472
 1546 GCGATGAAAGGCTCTTAACTGCGGTGATGTGTGATGATCAACGGTGTGCGCTTGTTTTG 1605
 473 GCTACGCTTCACTTCTTGAATGAGTCCGCTGCTGCTCTATGAGAGACGTGCTGTGTGTG 532
 1606 GGTATGCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
 533 TCGTCAAGTACCGGCTGAGAAATTTGTTTCTTCAACCAATGAGATCAGCATGCTCCGG 592
 1666 TCATCAGATACCGCTGAGT 1725
 593 GGAACCTGGGCTTCAAGACAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 652
 1726 GGAACCTGGGCTTCAAGACAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1785
 653 TCTTGGTGGGAGCCCAAGCTCTGTGACCATTTTGGAGTCCGCGGAGGACCATTAATG 712
 1786 ACTTTGAGGAGCAACCTGACCGTGTCAACATTTTGGCAGTCTGTGGGTGGAGCAAGT 1845
 713 TTTTGAAGTCTTAACTGTCTCCATGAGCAAAAGCTTATTCACAAAGCCATCATGAGAGA 772
 1846 TGTCTTCCGTTGT 1905
 773 GTGGGAGTGCATCATCCCTTACCTGAGAGCCCATGATTAAGAGAAAGATGAGAGACTG 832
 1906 GTGGGAGTGCATCATCCCTTACCTGAGAGCCCATGATTAAGAGAAAGATGAGAGACTG 1959
 833 AGGTGTTGACATTTCTGTGTGAACATGCGTCAAGACTCTGAGGCTGTGAGAGTGC 892
 1960 CGGTGTGTCACACCTGTCTGCTGTGACCAAGTGTGACTGTGAGGCTGTGAGTGC 2019
 893 TGAAGACAAACCTTCAAGAGTGTGACCTTCAAGCCCAAGAAACAAAGCTTTTCACTC 952
 2020 TCCGGGGCAAGAGTAAGAGAGATTTCTTGAATTAACAGACCTTTCAAGATGATCCCG 2079
 953 GAGTGGTTGATGAGTGTCTTTTCCAAAGAGCTTATGATGATGATGATGATGATGATGATGAT 1012

2080 GAGTGGTGAATGAGGCTTCTGCTGACAGGACCCCAAGAGCTGCTGCTGACCACT 2139
 1013 TTAAGCAATTTCTTCCATCATCGAGTGAATTAACAAGAGTGTGCTTCTGCTGCC-- 1070
 2140 TTACAGCTGTCTCTTGCATTTGTGTGTCAACAACATTAATTCGCTGCTCATCCCA 2199
 1071 -----TATGAAGAGGCTCTGAGATCTTCAGTGTGCTCAACAAGTCCCTTGCCC 1120
 2200 AGTCAATGAGATCTATGATTAACCAAGAGAAATGAGACAGAGGCTTCCAGGCTGCTC 2259
 1121 TCCATCTGATTAACAAACATCTGCAATCCGCTCAGATTTTGCACCTTGTGCTAATG 1180
 2260 TCAGAAATTTGTTAACGCTGTGATGTTGCTCTTCAATTTGATGACCTGAGAGGAG 2319
 1181 AATACTTCAATGACAAAGACTCCCTGACTGAAATCCGAGACAGTCTTCTGACTGCTG 1240
 2320 AGTACATTTGGGACAAATGGGATCCCAAGCCCTCAAGCGAGTTCCAGAGAGATGATG 2379
 1241 GAGATGTGTTCTTGTGTGCTCCCTGCACTGATCAAGCTGATATCAAGAGATGCTGTG 1300
 2380 CGAATCTCATGTTTGTGATTCCTGCACTCAAGTACATTTTCACTG---TTCCCGGG 2436
 1301 CACTGTCTACTTCTATGAGTTTCGGCAACCGCTCAGTCTTGTGAAGACAGAAACCGG 1360
 2437 CCGCTGTACTTCTTACAGAGTTCAAGCATCAAGCCAGCTGAGCTCAAGAACATCAGGCCAC 2496
 1361 CTTTGTCAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGCTGCTTCTGA 1420
 2497 CCGACATGAAGGACAGCCATGTTAA----- 2521
 1421 AGGGGACATTTTATGTTTGAAGGACCGAGGAGGAGAGATTACTGAGCCGAGAG 1480
 2522 -----ATTCACTGAGGAGAGAGAGAGCTTAAGCAGAGAA 2556
 1481 TGATGAATACTGAGGCTACTTGTGCTGCAACCGGAAATCTATGAGAAAGCATGCTGTC 1540
 2557 TGATGAAGTACTGAGGCAACTTTGGAGAAATGGGAAACCCATGCGAGGCTGTGCCAC 2616
 1541 TGTGGCAGCTTATTAATGACTGAGAGTACTTCACTGAGCTTGAACATGAGCTGTG 1600
 2617 ACTGGCCGCTGTTCACACGAGAGCAATATCTGAGCTGTAACCTACAGCCTGCGGTG 2676
 1601 GACAGAGACTCAAGAACCGGGGTGATTTTGAACGACCAATCCGCC 1651
 2677 GCGGAGCTTGAAGGCCACAGAGCTTCACTTCTGAGAGAGCGCTGCC 2727

RESULT 11
 HSM806270
 LOCUS HSM806270 3909 bp mRNA linear HTC 22-SEP-2004
 DEFINITION Homo sapiens mRNA; cDNA DKFZp686H0466 (from clone DKFZp686H0466).
 ACCESSION BX538086
 VERSION BX538086.1 GI:31874222
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 3909)
 REFERENCES
 Bloeker H., Boeher M., Brandt P., Mewes H.W., Well B., Amid C.,
 Oanger A., Fodor G., Han M. and Wiemann S.
 The German cDNA Consortium
 Direct Submission
 Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp686H0466) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.

Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/c1.cgi?cloneid=DKFZp686H0466>
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

Source

1. 3909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686H0466"
/db_xref="taxon:9606"
/clone="DKFZp686H0466"
/tissue_type="endometrium carcinoma cell line"
/clone_id="686 (synonym: hlcc3). Vector pSPORT1_Sfi, host
DH10B; sites SfiI + SfiIb"
/dev_stage="adult"
/note="carboxylesterase 2 isoform 2"
1. 3909
/gene="DKFZp686H0466"
1002. 2825
/gene="DKFZp686H0466"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD98009.1"
/db_xref="GI:31874223"
/db_xref="UniProt/Swiss-Prot:O00748"
/translation="MTAQSRSPTTFPPSPSPTPLTPCPVOTPRGLKALHICMTDPG
OPLGROQRVRORTSEPTSEPTMLRLRLSLVAGSLILVRRGGDSASIPRTHT
GOVLGLVHKNAGNVOTFLGIIPAKPIGLRPARPPSPSNGVSDGTTHPMCL
ODLTVESFRLSQPNMTFSPSDMSDECLYFTSTYPMHSGNSLIVMVAHGLAVFG
NASLIDGSMALLENVVVITQIRGLGVLFSTYGDHAGNNGVLDVVAALWVONI
AHFVGNPDRVTI FGSAGTSVSLVSPISQGLPHGLAMEGVALLPGLIASADVI
STVVANLSACDDVDSBALGVCLRSKESLILAIINKPKMIPEVVDGFLPRPDLA
SADPOVPSIIVGVNNNEFGMLIPKVMRIYDTOKEMDRASQAALIKMLTLMLPTSG
DLIREYIGDNGDPOTLOAFQEMMAWSFVTPALOVHFOCSRAPVYFLEPHQPSM
LKNIRPPHKKADHVKTREERDLSRKMKVMANPRANKPNSEGLPHMPLFQESQYL
RLNLOPVRKRLKAHRLQFMKALPQKIQELSEPERHTEL"

gene

CDS

Query Match 19.7%; Score 344.2; DB 3; Length 3909;
Best Local Similarity 54.7%; Pred. No. 1.1e-86;
Matches 849; Conservative 0; Mismatches 618; Indels 84; Gaps 5;
ORIGIN

116 GGAACACGAGCTGGATGATTCAGGGGCAAGCAATGCTGTGTGGAAAGCCCTGTGC 175
Db GGACACACACACGCGGGGAGGTGTGGGGAAGCTTGTCCATGTGAAGGCGGCAATGCG 1353
QY CTGTGAAGTGTCTCGGAGTCCCTTGTGTGCTCCCGCTGGGATCCCTGGCATTTA 235
Db GGATTCAAACCTTCTGTGGAATTCATTTGCAAGCCACCTCTAGTCCGCTGCATTGG 1413
QY CGAACCCGACGCTGCATCGCCCTGGGATTAATTGCGAAGCACTCTACCTTAATT 295
Db CACCCCTGAGCCCTGAACTTGTGAGTGGGTGAGGATGGAACCAACCATCCGCGCA 1473
QY TGTGCTTCAAGAACTCAAGTGGCTCTTGA---TCAACACATGCTCAAGTGCATT 352
Db TGTGTCTACAGAGACTCAACCGCAGTGAAGTCTTGAACAGTTCAACATGACCT 1533
QY ACCGGAATTCGAGTGTCAAGAACCTGCTACCTGAACATCATATAGCGCTCCACG 412
Db TCCCTTCGATCCATGTCTAGAGATGCTGTACCTTCAAGCATTCACGCGGCCCATTA 1593
QY CCGATTAACGCTCAAGCTCCCGCTTGTGTGTGTGTCCACGAGAGTCCCTTCAAGACTG 472
Db GCCATGAAGGCTTAACCTGCGGATGATGTGTGATCAAGGTGTGCGCTGTGTTTTG 1653
QY GCTAAGCTCATTTTGAATGGGTCGCGCTGGCTGCTTATGAGAGAGTGTGTTTGG 532
Db GCATGAGCTTCTTGTATGATGTTTCATGTGCTGCTTGGAAAGGTGTGTGTGTCA 1713
QY TCGTCAAGTACCGGCTAGAAATATTGGTTTCTTCAACAATGGGATCAGATGCTCGG 592
Db TCATTCAGTACCGGCTGGGTCTCTGGGCTTCTTCAAGCATGAGAGCAAGCAACG 1773

QY 593 GAAACTGGGCTTCAAGACACAGTGGCTGTCTGTCTGGGTCCAGAAATCATGAGT 652
Db 1774 GCAACTGGGGCTACTGACCAAGTGGCTGCACTACGCTGGGTCCAGAAATATGCCCC 1833
QY 653 TCTTGGTGGGACCCCGAGCTGTGACCATCTTTGGGAGTCCCGGGAGCCATTAAGT 712
Db 1834 ACTTGGAGGCAACCTGACCGGTGACCATTTTGGGACAGTCTCGGGTGCAGAGTG 1893
QY 713 TTTAGTCTTATACGTCTCCATAGCCCAAGGCTTATTCAGAAAGCATATGAGA 772
Db 1894 TGTCTTCGTTTGTGTGTCCTCCCATATCCAGAGACTTTCACGAGACATATGAGA 1953
QY 773 GTGGGGTCCATCATCCCTTAACGTGAGGCCCATGATTAAGAAAGTGAAGCATGCG 832
Db 1954 GTGGGTGGCTCTCTGCGCGGCTCATATGTCAGCTCAGCTGA-----TGTCACTCCA 2007
QY 833 AGTGGTTGCATTTCTGTGTACATATGCTCAGATCTTGAAGCCCTGCTGAGTGC 892
Db 2008 CGGTGTGGCCAACTGTCTGTGTGACCAAGTTGACTTGAAGGCCCTGTGGGCTGCC 2067
QY 893 TGAGACAAACCCCTCAAGAGCTGTGACCCCTCAGCCAGAAAACAAAGTCTTTCATCTC 952
Db 2068 TCGGGGGCAAGATTAAGAGATTTCTGCAATTAACAAGCTTTCAGATGATCCCG 2127
QY 953 GAGTGTGATGTGCTTCTTCTTAATGAGCTTGAATCTATTTGTTCAGAAAGCAT 1012
Db 2128 GAGTGTGATGTGGGTCTTCTCTGCGCCAGGCAACCCCAAGAGACTGCTGCTGCCACT 2187
QY 1013 TTAAGCAATCTCTTCATCATCGGAGTCATTAACGAGTGTGGCTCTGCTGCC-- 1070
Db 2188 TTCAAGCTGTCCCTAGCATTTGTGTGTCACCAACATTAATGCGGTGCTCATCCCA 2247
QY 1071 -----TATGAAGAGGCTCTGAGATCTCAGTGCTCCAGAAAGTCCCTGCC 1120
Db 2248 AGGTCAAGAGATCTATGATACCAAGAAATGACAGAGAGGCTCCAGGCTGCTC 2307
QY 1121 TCCATCTGATCAAAAACATCTGCAATCCGCTCAGTATTTGCACTTGTGGCTAATG 1180
Db 2308 TGCAGAAAATGTTAACCTGTGTATGTGCTCTTACATCTGTGACCTGTGAGGAGG 2367
QY 1181 AATACCTTCATGACACACATCCCTGACTGAATGGAATCCGAGACAGTCTTGGACTTGTGCTG 1240
Db 2368 AGTACATTTGGGACAAATGGGATATCCAGACCTTCAGAGCACTTCAAGAGATGATG 2427
QY 1241 GAGATGTGTTCTTGTGTGCTCTGACTGATCAAGCTGATATACAGAGATGCTGTG 1300
Db 2428 CGGACTCCATGTTGTGATCCTGCACTCCAGATGACATTTTCAAGT---TTCCCGGG 2484
QY 1301 CACCTGTCTACTTATGAGATTTCCGACCGGCTCAGTGTCTTGAAGACAGAGCCGG 1360
Db 2485 CCGCTGTACTTCTAAGAGATTCAGATCAAGCTCAGCTGCTCAAGAAATCAGGCGAC 2544
QY 1361 CTTTGTCAAGCGGACCAAGCTGATGAAGTCCGCTTGTGTGCGGTGCTTCTGTA 1420
Db 2545 CGCAATGAAGGAGACATGTTA-----ATTCATGAGGAAGGAGCAGCTTAAGCAGGAAGA 2569
QY 1421 AGGGGACATTTGTATGTTGAAAGAGCCACGAGAGAGAAATTACTGAGCGGAGAGA 1480
Db 2570 -----ATTCATGAGGAAGGAGCAGCTTAAGCAGGAAGA 2604
QY 1481 TGATGAATATCTGGGCTACCTTGTGCGAACCGGGAATCTTAATGGAAACGACCTGTCTC 1540
Db 2605 TGATGAATATCTGGGCTACCTTGTGCGAAATGGAAACCCCAATGGCGAGGCTTCCAC 2664
QY 1541 TGTGCGAGCTTAATTAATGATGACAGATCACTCCAGCTGATGAAATGAGAGCTCG 1600
Db 2665 ACTGCGCGCTGTGACCAAGAGAGCAATACCTGTGGCTGAACCTAAGCTGCGGTGG 2724
QY 1601 GACAGAGATCAAAAGAACCGCGGTGATTTTGAACGACACATCCCTC 1651
Db 2725 GCGGGCTCTGAAGGCCCAAGGCTTCAGTGTGGAAGAGCGGCTGCC 2775

LOCUS	CP111083	1918 bp	mRNA	linear	EST 23-JUL-2003
DEFINITION	Shultzomic04334 Rat lung airway and parenchyma cDNA libraries				
ACCESSION	Rattus norvegicus CDNA clone Contig3845 5', mRNA sequence.				
VERSION	CP111083				
KEYWORDS	CP111083.1	GI:33167595			
SOURCE	EST.				
ORGANISM	Rattus norvegicus (Norway rat)				
REFERENCE	Rattus norvegicus				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
TITLE	1 (bases 1 to 1918)				
JOURNAL	Shultz, M.A., Zhang, L., Gu, Y.-Z., Baker, G.L., Fannuchi, M.V., Padua, A.M., Gurske, W.A., Morth, D., Penn, S.G., Jovanovich, S.B., Plopper, C.G. and Buckpitt, A.R.				
COMMENT	Gene expression analysis in response to lung toxicants: I. Sequencing and microarray development				
	Am. J. Respir. Cell Mol. Biol. 30 (3), 296-310 (2004)				
	Contact: Shultz MA				
	Dept. of Molecular Biosciences, School of Veterinary Medicine				
	University of California, Davis				
	1311 Haring Hall, One Shields Avenue, Davis, CA 95616, USA				
	Tel: 530 752 0793				
	Fax: 530 752 4698				
	Email: maeshultz@ucdavis.edu				
	Average Phred score is 20 or better. All poor quality data (Phred < 20) and vector/linker sequence has been removed.				
FEATURES	High quality sequence stop: 1918.				
source	location/Qualifiers				
	1..1918				
	/organism="Rattus norvegicus"				
	/mol_type="mRNA"				
	/strain="Sprague-Dawley"				
	/db_xref="taxon:10116"				
	/clone="Contig3845"				
	/sex="male"				
	/tissue_type="airway or parenchyma"				
	/der_stage="adult"				
	/clone_lib="Rat lung airway and parenchyma cDNA libraries"				
	/note="Organ: lung; Vector: pGEM-11ZF(-); Site: 1: Eco RI; Site 2: Not I; mRNA was isolated from microdissected rat lung airways and parenchyma tissues."				
ORIGIN					
Query Match	19.4%; Score 338; DB 7; Length 1918;				
Best Local Similarity	54.2%; Pred. No. 5.5e-85;				
Matches	869; Conservative 0; Mismatches 660; Indels 75; Gaps 6;				
Gy	91	GGGCTTGTGTGAAGGGCCACAGAGAAACACAGGCTGGAGTATCAGGCGAAGAA	150		
Db	83	GGGTATTCATCTCTCACCACTGTGTGAACACTGTAAAGCAAGTCTTGGGAAGAT	142		
Gy	151	GTCACGTGTGTGGGAAGCCCTGTGTCTGTGAACGTTCTCTGGAGTCCCTTGTGCT	210		
Db	143	GTCATTTTGGAGAAGATTGGACAGGCTGTGGCTGTCTTCTTGGGAATCCCTTGGC	202		
Gy	211	CCCCCGCTGGATTCCTTCGATTACGAATCCGACGCTTGATGCTCCCTGGATAC	270		
Db	203	TCCCTCTTGAGCTCTTGAGGTTTGCTCACACACAGCTCGAGAGTCTTGGAACT	262		
Gy	271	CGAGAAGCACTCTCTACCTTAATTGNGCCTCCGAATC-----AGATGGCTGCT	324		
Db	263	AAGATATCTACTCTCTTACCACTATGTGCTCTCAAGTGTCTTGGAGGGCAGTTCT	322		
Gy	325	TTAGATCAACACATGCTCAAGGTGATTAACCGAATTCGAGTGTCAAGAGATGCTCT	384		
Db	323	TCAGAGCTTTTACCAACAGGAAGAAACATTCCTTACAGTTTCTGGAAGATGCTCT	382		
Gy	385	TACTGAAACATGTATGCGCTGCGCAGGCGAATACAGGCTCAAGCTCCCGTCTGTG	444		

Db	383	TACCTGAATGTTTAACTACCCGGCTGACTTGACAAGAAAGACGGCGTTACCAGATGATGGTG	442
Oy	445	TGTTTTCCAGAGAGTGCCCTTCAMAGCTGCTCAGCCTTCATCTTTGATGGGTCCGCGCTG	504
Db	443	TGATTCATGAGAGGTGACTGTGATGGGTGGATCATCTCATATGATGACAGAGCCCTC	502
Oy	505	GCTGCCATATGAGGACGTGCTGGTTTGGTGTGTCCAGTACCGGCGCTAGGAATATTTGGTTTC	564
Db	503	TCCTCCCAATGAAAATGTGTGTGTGTGTACATTCATGATATGCGCTTGGCATCTGGGGATTC	562
Oy	565	TTCAACCATATGAGATCAGCATGTCTCCGGGGAACCTGGCCTTCAAGGACCAAGTGGCTGT	624
Db	563	TTCAAGCAGAGGAGATGAAACAAGCAGGAGGACACTGGGGTCAATTTGGACAGAGCTGGCTGA	622
Oy	625	CTGTCTGGGTCCAGAAAGAACATCGAATTCTTGGTGGGACCCAGCTCTGTGACATC	684
Db	623	CTGCATCGGGTCCAGACAATATGCGCAATTTGGGGGCAACCCAGGCTGTGTACATC	682
Oy	685	TTTGGCGAGTCCCGGGGAGCCATPAAGTTTCTAGTCTTATPACGTCTCCATGGCCAAA	744
Db	683	TTTGGAAATCTGAGAGGTTTCAAGTCTCTGTTTGTGTATCTCTTGGCCAAAG	742
Oy	745	GGCTTATTTCCAAGAGCATCATGAGAGTGGGTGGCCATCATCCCTTACCTTGAGAGCC	804
Db	743	AACCTCTTCCAGAGGCGCATTTCTGAGAGTGTGTCTCTCAATATCGCTGTGATTTACA	802
Oy	805	CATGATTAATGAGAAAGTGAAGACCTGCAAGTGTGTCACATTTCTGTGTACATATGC	864
Db	803	ACAGATGACAAAGCCCATTTGCTAATCTGATTTGCTACTCTTCTGGGTGTAACAACACACA	862
Oy	865	TCACACTCTGAGGCGCTGCTGAGGTGTGCTAGACA-----	900
Db	863	TCAGCTGTATATGTTCAATGCTGCGCCAGAAAGACAGAGATGAATCTCTGAGACATTCA	922
Oy	901	-----AAACCTCCAGAGAGCTCTGACCTCAGCCAGAGAAACAAAGTCT-----	945
Db	923	TTAAATTTGAATCTTTTAACTGAGACTTATTTGAAACCCAAAAGAGGTATCCCTTC	982
Oy	946	TTCACTCGAGTGTGATGTGCTTTCTTTCCTAATAGAGCCTCTAGATCTATATGTCTCAG	1005
Db	983	CTACTCTATGTGATTTGATGAGTGTGCTGTCAAAAGACACAGAAAGATCTCTGGCTGAG	1042
Oy	1006	AAAGCATTTAAAGCAATTCCTTCCATCATCGGAGTCAAATAACACAGATGTGCTTCTGT	1065
Db	1043	AAAGATTTCACAAGAGTCCCTCATATGTGTGGCATCAACAAGCAAGATTGGCTGATC	1102
Oy	1066	CTGCGCTATG-----AAGAGGCGCTCGAGATCCTCAGTGGCTTCAACAAGTCCCT	1115
Db	1103	ATTCCAATGTCTTATGGGCTATCCACTTTCGAGAGGCAATCTGACACGAAAACAGCCAAAT	1162
Oy	1116	TGCCCTCATCTGTATCAAAAC--ATCTGACATCCCGCTCAGTATTTTGACACTTGTG	1173
Db	1163	TCCTCTTTGTGGAAGTCTTACCCAAACACTGAAATCTCTGAGAAAATGATTTCCAATGGCT	1222
Oy	1174	GCTAATGATATCTTCCATGACAGCACTCCCTGACTGAATCCGAGCAGTCTTCTGAC	1223
Db	1223	TCTGAGAAAGTATTTAGAGGGGACAGATGACCTTCCGCAAAAGAAAGACTGTTCAMAGAC	1282
Oy	1234	TTGCTGTGAGATGTGTTCTTTGTGTGCTCCGCACTGATTCACAGCTCGATATCAAGAGAT	1293
Db	1283	TTGGTTGCAATGATGATATTTTGTGTCTCCATCAATGATGTGTTCGAAGTCAAGAGAT	1342
Oy	1294	GCTGTGTGACCTGTCTACTTCTATGAGTTTCGACACGGCGCTCAGTCTTTGAAGACAGC	1353
Db	1343	GCTGGAAGCCCACTTCATGATGAAATTTGATGATGCGCCCAAGCTTTGTATCAGCATG	1402
Oy	1354	AAGCGGCTTTTGTCAAAGCCGACACAGCTGATTAAGTCCGCTTTGTGTTCGTGTGCGC	1413
Db	1403	AGGCGCAAGACAGTGAATCGAGACCAATGATGATTAATCTTCTCAGTATTTGATCTCCA	1462
Oy	1414	TTCCGGAAGGGGGAATTTGATTTGTGAAAGAGCCAGGAGGAGGAAGTTACTAGC	1473
Db	1463	TTTTTAAH-----AGATGGCGCTTCAAGAGGAGACCAATCTCAGC	1504

QY 1474 CGAAGAGATGAATTAATGGCGTACCTTTTGTCGAAACCGGGAATCTTAATGGGAACGAC 1533

Db 1505 AAAATGGTGAATGAATTAATCTGGGCCAACTTTGTGTGGAAATGGGAACCTTAATGGGGAGAGG 1564

QY 1534 CTGTCTCTGTGTGGCCAGCTTAAATATATGTGACTGAGCAGTACTCTCAAGCTGGAATTGAACATG 1593

Db 1565 CTGGCCCACTTTGGCCGAATATATGACCAAGAGGAAGGATCTCGAAGATTGTGTCTCAACT 1624

QY 1594 AGCCTCGAGCAGAGACTCAAGAAACCGGCGGTGATTTTTGGAC 1637

Db 1625 CAGCAGAGCCAGAGGCTGAAAGCAAAAGAAATGGCTTTTGTGTC 1668

RESULT	13
LOCUS	BC019926
DEFINITION	BC019926 2038 bp
ACCESSION	Mus musculus, clone IMAGE:5123923, mRNA.
VERSION	BC019926 GI:18044766
KEYWORDS	HTC.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
	linear HTC 20-SEP-2002

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (baaes 1 to 2038)
Strausberg, R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health. Mammalian

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov	Contact: MGC help desk

Tissue Procurement: Jeffery R. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kovlis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clome distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINt at: <http://image.ltn.gov>
 Series: IRK Plate: 42 Row: D Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarly but not identical to protein
 This clone has the following problem: frame shifted.

FEATURES	Location/Qualifiers
SOURCE	1. .2038

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:512323"
/issue_type="Liver, normal, 5 month old male mouse."
/clone_id="NCI_CGAP_l19"
/lab_host="DH10B"
/notes="Vector: pCMV-SORT6"

```

ORIGIN

Query Match	19.0%	Score 331.4;	DB 3;	Length 2038;
Best Local Similarity	53.1%	Pred. No. 4.4e-83;		
Matches 839;	Conservative	0;	Mismatches 711;	Indels 31; Gaps 5

88 ACTGGGGCTTCTGCTGAGGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAG 147

Db	98	ACTGACCCAAAGTCATTCAGCCTGGAAGTGGATACCCCTGGGTCTGTTCCAGGCCG	15
Qy	148	CAAGTCATGTGCTGGAGAGCCCTGTGCTGTGAAOGTGTTCCTCGAGTCCCTTGTCT	207
Db	158	CAGGTGGGTGTGAAGACACAGACCCCATGGTAAATGTCTTCTCGGCAATCCCTTGTCT	21.7
Qy	208	GCTCCCGCTGGGAATCCCTGGCATTTTATGAAACCCGACGTGATACCGCTTGGATAC	267
Db	218	CAGCACACAGTGGACCTCTTCGGTTCTACGCTCACCTCCACACACAGCCTGGAAAGT	277
Qy	268	TTGCGAAGAACCACTCCCTAACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTCTTA	327
Db	278	GTEGAGAGATGCCAGATCAATCCCCCAATGTGCTTCAAGATGTAGAAATAATCAAC	337
Qy	328	GATCAACACATGCTCAAGTGTGCTTACCCGAAATTGGAGTGTCAAGAATCTGCTTAC	387
Db	338	AGCAGATTTGGCTTCATATAAAAGATTAATAATCTTCCCATTTCTAGAGACTGCCAGAC	397
Qy	388	CTGAACATCTATGCGCTCTCCACGCGGATACAGGCTCCAGAGCTCCCGCTTGGTGG	447
Db	398	CTCAACATCTTACGCCCCCATGATATCTGCAGGGGACAAAGCGGCTCATGTATGG	457
Qy	448	TTCCACAGAGTGCCTTCAAGACTGGCTCAGCCTCATTTTGAATGGTCCGCTGCT	507
Db	458	ATCCACGAGGCTCTCTGCTGGTGGCTCTCCACATCTCAGAGATGATCAGACTGGCC	517
Qy	508	GCCTATAGAGACATGTGCTGGTTGGTTCGTCAGATCCGCGTACGAATATTGGTTCTTC	567
Db	518	GGCTACGGGGAATGTGTATGTTATCATCTGTCAATATGCTTGGGATCTTGGCTTCTC	577
Qy	568	ACCACATGGGATCAGCATGCTCCGAGGAACTGGGCTTTCAGAGACAGGTGGCTGCTG	627
Db	578	AGCATTGAGACAAACACATGCCAGGCAACAGGGATTCCTGGATGTGGTGGCTGCTT	637
Qy	628	TGCTGGGTCCAGAAACATGAGTTCTTCGGTGGG-GACCCAGCTCTGTACCATCTT	686
Db	638	CGCTGGGTCCAGGGAACTAGCCCCCTTTGGGGGTATATCCCACTGTGTCACTATCTT	697
Qy	687	TGGCAGTCCGCGGAGGCATTAATGTGTTCTTATGTTTACTGTCTCCCATGGCCAAAG	746
Db	698	CGGTAACTCTGTGGAGGCATGATTTGTCTCATCCCTGTTCTGTCTCCAAATATCTGCTGG	757
Qy	747	CTTATTCACAAAGCATCATGAGAGTGGGGTGGCCATCATCCCTTACTCTGGAGGCCCA	806
Db	758	GCTCTTCCACAGAGCATATCCACAGTGGGATTTGTCAAGCACTTATGATGATG-----A	811
Qy	807	TGATTTATGAGAAAGTAGAGACCTGCAGGTGGTTGGCAATTTCTGTGGTAAACAATGCTTC	866
Db	812	AGACATGAAACATGTGCTGTAAAGCTCAGAACTTTGCCAATTTCTGTGGCTGTGGCTCTGC	871
Qy	867	AGACTCTGAGGCCCTGTGCTGAGGTGCTGAGGACAAACCTCCAAAGAGTGTGACCCCT	926
Db	872	ATCCCCAGCTGAGCTGGGTCCAGTGTGTTGCTCAGAAAGAGAAAGACCTT-----AT	925
Qy	927	CAGCCAGAAACAAAGTCTTTCATCTCGAGTGGTTGATGTGTCTTTCTTCTAATGAGCC	986
Db	926	CAGCAGAAACAGTGAACATTTCTTCAATATGATCAATGATCTCTTCTTCCACAAAGGCC	985
Qy	987	TCTGATCTATTTGCTCAGAAAGCATTTAAAGCAATTCCTCCATCATCGAGTCAATTA	1046
Db	986	AGAGAGTCTCTAGCAGACGAGCAATTCGCCACGTGTGCCCTACTTTTGGAGTCAACCA	1045
Qy	1047	CCACGATGTGCTTCTGTC-----TGCCTATGAAGAGCTCCTGAGATCTT	1094
Db	1046	CCATGAATTTGGCTGGCTTCTTCACTCAATCCTTGAAATATCCTGGATTAAGTTGAACATCT	1105
Qy	1095	CAGTGGCTCCAAAGTCCCTTGGCCCTCAATCTGATACAAACAACTCTGCACATCCGCC	1154
Db	1106	GAGCCGGGAAAGCTGTGTGAGATTTTACAGGCCCTTTTGGCCATTAATGGAAGTGGCCC	1165
Qy	1155	TCAGTATTTGACCTTGTGGCTATGATATCTTCATGACAAAGCATCTCTGACTGAAT	1214
Db	1166	TGAGATCATCCCACTGTCAATCATGATGATATCTCAGACAAATGGGCTCAATCAATGACTAC	1225

QY 1215 CCGAGACAGTCTTCTGAGCTTGAGATGATGTTCTTGTGTCCTCCAGCATGATCAC 1274
DB 1226 AAGGATGCTTCAGAGATTCCTGGATATCTCTTCATCATCTTCACTTGAACTT 1285
QY 1275 AGCTGATATCAAGAGATGCTGTGACCTGTCTAATTATGAGTTTGGCACCAGCC 1334
DB 1286 CTCAAAATACCTTCAGATGCTGGGTCCTGTCTTCTGTAGAGTTCCAGCATACACC 1345
QY 1335 TCAGTCTTTGAACACGAAACCCGCTTTTTCAAAGCCGACAGCTGATGAAGTCCG 1394
DB 1346 CAGTCTTTTGAAGATTCAGACCCAGCTGGGTGAAGCTGACATGCTCTGAGAAATTC 1405
QY 1395 CTTTGTGTCGGTGTGCTCTTCTGA-----AGGGGACATTTGATTTGGAAGAGC 1448
DB 1406 CTTTGTGTCAGAGCTCTTCTCTCACTGATGAGATTCCTTCTGGCTTCCAGAGCC 1465
QY 1449 CACGAGAGAGAGAAATTAATGAGCCGGAAGATGATGAATAATGAGGCTAATCTTGTCTG 1508
DB 1466 CACAGAGAGAGAGAGAGCTGAGCTGACATGATGAGCCCAATGAGCCAGTTTGACAG 1525
QY 1509 AACCGGGAATCTTAATGAGAGAGCTGTCTGTGAGCAGCTTAAATCTGATGAGCA 1568
DB 1526 CACAGAGAAACCCCAATGCGAGAGGCTGTCTTGGCCCAATTAACCACTTGAAGACA 1585
QY 1569 GTACTCCAGCTGAGCTTGAATGATGAGCTCGAGACAGACTCAAGAAACCGGGGTGA 1628
DB 1586 ATACTGGAGATTTGCTTGAATCAAGCACTGGGGTGAAGCTTAAGAGAGGTGGCTACA 1645
QY 1629 TTTTGTGACACAGACCATCCC 1649
DB 1646 GTTCTGACAGAGACACTGCC 1666

RESULT 14
LOCUS CR618303 1826 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSDB014YC12 of Placenta of Homo sapiens
(human).

ACCESSION CR618303.1 GI:50499110
VERSION HTc; CnSLT cDNA.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact: Peng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1826)

REFERENCE
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
source location/Qualifiers

1..1826
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDB014YC12"
/issue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 17.1%; Score 298.4; DB 3; Length 1826;
Best Local Similarity 57.1%; Pred. No.1.2e-73;
Matches 564; Conservative 0; Mismatches 422; Indels 3; Gaps 1;

QY 88 ACTGGGCTTCTGCTGAAGAGGCGCAGAGAGAACACCAAGCTGGATGATTCAGAGGCAAG 147
DB 84 ACTGGGCGCAAGATTGCTCAGCTGATGATGACACACACCTGGTCTGCTGCGAGGCGCG 143
QY 148 CAAGTCACTGTCTGGGAAGCCCTGCTGCTGGAAGGTTCCTGGAGTCCCTTTGCT 207
DB 144 CAGGTGGGCTGAAGAGGACAGACCGCTTGTGAATGCTTCTGGGCAATTCATTGCGC 203
QY 208 GCTCCCGCTGGATATCCCTGCAATTTACGAACCCGACGCTGATCCGCTGGATTAAC 267
DB 204 CAGCGGCACTGGGCGCTGACCGGTTCTCAGGCCCAACACCAAGCAAGAGCCCTGGAGGCT 263
QY 268 TTGGAGAAAGCCACTCTTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTTA 327
DB 264 GTGGGGAATGCCAGCATGCGCCCAATGTGCTTAAGAGAGTGAAGCATGAACAGC 323
QY 328 GATCAACATGCTCAAGTGTATACCGAAATTCGAGTGTCAAGAAATCTGCTCTTAC 387
DB 324 AGCAGATTTGTCTCAAGGAAACAGAGATCTTCTCGTTCAAGAGATGCTGCTGTG 383
QY 388 CTGAACATCTATGGGCTGCGCCAGCGCATACAGGCTCAAGCTCCCGCTTGTGTTGG 447
DB 384 CTCAAGCTCTATAGGCCAGAGCTGAGTCCCGCAGAGGTCCGATGAGCCGCTCATGATAG 443
QY 448 TTCCAGAGAGGCTCTTCAAGACTGCTCAGCTCATCTTGTGATGGTCCGCTGCT 507
DB 444 GTCCATGAGAGGCGCTCTGATTAATGGCGTGCACCTCTCAAGATGATCAGCTGTGGT 503
QY 508 GCTTATGAGAGCTGCTGCTGTGTGCTCCAGTACCGGCTAAGAAATTTGGTTCTTC 567
DB 504 GCTTATGAGAGTGTGTGTGTGCTTACAGTCCAGTACCGCTTGTGGCTCTTTC 563
QY 568 ACCAATGAGATCAGACATGCTCCGGGAACTGGGCTTCAAGAGACAGGTGCTGCTG 627
DB 564 AGCACTGAGAGATCAGATGACCTGGCAACAGAGCTCTTGAATGTGATGCTGCTTTG 623
QY 628 TCTGGGTCCAGAGAAATCAGAGTCTTCTGCTGGAGAACCCAGCTCTGTGACCATCTTT 687
DB 624 CGCTGGGTCAAGAAATCATGCGCCCTTGGGGGTGAATCACTGTGTCACTGCTTT 683
QY 688 GCGAGTCCGCGGAGCCATTAATGTTTCTTACTTATCTGTCTCCATGAGCCAAAGGC 747
DB 684 GGTGATCTGCGGAGGAGCATCTCTGTGCTGTGCTGTCTCCAGTGTGCTGAGGG 743
QY 748 TTAATCCAAAGCATATGAGAGTGGGGGCGCATATCCCTTACCTGAGGCGCAT 807
DB 744 CTGTTCAAGAGCCATCAACAGAGTGGGTGATCAACACCCAGGATCATCACTCT 803
QY 808 GATTATGAGAAAGTGAAGACTGAGAGTGTGACATTTCTGTGTAACATGCTCA 867
DB 804 CACCTCT---TGCCCTTACGCTCAAGAAATCGAAACACCTTGCGCTGACGCTCC 860
QY 868 GACTGTGAGCGCTCTGCTAAGTGTCTGAGACAAACCTCCAGAGAGCTGTGACCTTC 927
DB 864 CCGGCTGAGATGTGCACTGCTTCAAGCAAGAAAGAGAGAGACTGCTCTTGTGCAAG 920
QY 928 AGCCAGAAACAAATCTTCACTGAGATGTGTAATGAGCTTTTCTTAATAGAGCT 987
DB 921 AAGCTGAATAATCTATCTCTCTCAACGTTGATGAGATGCTTCCCAAAAGCCCC 980
QY 988 CTAGATCTATTGTCTCAAGAAAGCATTTAAAGCAATCTTCATCATGAGAGTCAATAC 1047
DB 981 AAGGAATCTTGAAGAGAGAGAGCCCTTCACTGTGTGCTTCTCAATGGGTGTCAACAC 1040
QY 1048 CACGAGTGTGCTTCTGCTGCTATGA 1075
DB 1041 CATGAGTTCAAGCTGCTCATCCCCAGGA 1068

```

RESULT 15
CR641368      1859 bp      mRNA      linear      HTC 18-AUG-2004
LOCUS      CR641368
DEFINITION      Tetraodon nigroviridis full-length cDNA.
ACCESSION      CR641368
VERSION      CR641368.1 GI:51137813
KEYWORDS      HTC; cDNA; full-length; Tetraodon nigroviridis.
SOURCE      Tetraodon nigroviridis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Acanthopterygii; Neopterygii; Teleostei; Euteleostei;
      Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
      Tetraodontidae; Tetraodontidae; Tetraodon.
      1 (bases 1 to 1859)
REFERENCE      Genoscope.
AUTHORS      Direct Submission
TITLE      Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage -
      : 2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
      (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
      The sequences are based on single pass reads.
      More information available at
      http://www.genoscope.cns.fr/tetraodon.
FEATURES
     source      1..1859
                  location/Qualifiers
                  /organism="Tetraodon nigroviridis"
                  /mol_type="mRNA"
                  /db_xref="taxon:99883"
                  /label_type="Liver"
ORIGIN
Query Match      16.8%; Score 293.6; DB 3; Length 1859;
Best Local Similarity 51.9%; Pred. No. 3e-72;
Matches 771; Conservative 0; Mismatches 694; Indels 21; Gaps 4;

104 AAGGCCACAGAGAAACACGAGTGGAGTGGATTGAGGCAAGACGACCTGCTGG
108 ATGCACCTGAAATCCATACCAACCTCGGAGCTGAGAGGTGAATATCAGACCGTGAAGG
164 GAAGCCCTGTGCTGTGAGACGCTGCTCGAGTCCCTCTTGTCTGCTCCCGCTGGG--
168 GGAAGAAGCGTGGAGTCCATGATACCTGGGTGCTCCCTTGTCTAAGCACTCTGGGCT
222 - ATCCCTCGCATTTGGAACCGCGACGCTGCATGCGCCCTGGGATTACTTGGAGAACCA
228 CTGCTGAGGCTGCTCACCCTCGCTGTAGAGGCTTGGGAAGAGTAGAGATGCTA
281 CCTCTACCTTAATTGTGCGCTCCAGAACTCAGAGTGGCTGCTTATATCAACATGC
288 CCAAGCAACCGCTGATGTGTGTTTCAGATTTAGAGTTGCAACACAAATTTAAAGAAC
341 TCAAGGTGCATTAACCCGAATTCGAGTGTGAGAACTGCTCTACCTGAACATCTATG
348 TTAATTTAGAGTCAACCTTCAGACATTTCAAGAACTGCTTACTCAACATTTTACA
401 CGCTGCCCAAGCCGATACAGGCTTCAAGCTCCCGCTTGTGGTGTGTTCCAGAGGTG
408 CTCTGCCAACAAGCTGAGAAATACCAAGCTCCCAAGTATGTCTGATCCACGGTGAAG
461 CCTTCAAGACTGCTCAAGCTCCATCTTTGATGGGTCCGCTGGCTGGCTGATAGAGAG
468 GGTGTGTTTAAAGTTCGGCTTCAATGTATAGTGGCTTGGCTTGGCTTCAACAAATG
521 TGTGTTGTGTGTGTCGCAAGTACCGGCTAGGAATATTGTTTCTTCAACCATGGATC
528 TTGTTGTGTGTGTATCCAGTACCGACTGGGAATACGGGCTTTCTGAGCACTGTGTATA
581 AGAATGCTCCGGGGAACTGGGCTTCAAGAACAGTGGCTGCTGTCTGCTGGSTCCAGA
588 AGCATGATCCAGGAACAATGTGTTCTGTGATCAGGTCCAGGCTTGAATGGGTCCAGG
641 AGAATGATGATTTCTGTGTGGGAGCCCAAGCTCTGTGACCATCTTGGGAGTCCGGCG

```

```

Db      648 AGCATATTCAACAATTGGAGAGAACCCAGATTATATGACCAATTTGGCGAGTCTGTG
QY      701 GAGCATTAAGTGTTCCTAGCTTTATATCTGTCTCCATGGCCAAAGCTTATTCACAAAG
Db      708 GTGAGTGTAGTATCTCTGCTGCTCAACCTGTGTGAGGGCTTTTCCGCCACG
QY      761 CCATCATGAGAGTGGGGGCGCATCATCCCTTACTGAGGCCATGATTATGGAAGA
Db      768 CAATGTCTGAAGGTGCACTGCTGCAATGATATATTTG-----CTAAGATGACCCCC
QY      821 GTGAGACCTGCAAGGTGTGTCACATTTCTGTGTATCAATGCGTCAGACTGTGAGGCC
Db      822 ACCCATGTATACATTTGTGCGAAATGATCTGTGTGATGACATGAACACAGAAAAA
QY      881 TGTGAGGTGCTGTGAGAACAAAACCTTCAGAGAGCTGTGACCTTCAGCCAGAAAAA
Db      882 CTGCTGATTCATCAAGAACTGTGACGTGAGAAACATTTTACCTTATCTAAGAAAAAG
QY      941 AGCTTTTCACTGAGTGTGATAGTGTCTTCTTCTTAATGAGCTCTAGATCTATGT
Db      942 AGTTAAGTATCTTTAATCAACCGATGATATTTCTTCAAAAAATCAGCAGAAAGTTGT
QY      1001 CTC--AGAAACATTTAAAGCAATTCCTTCATCATGAGTCAATTAACACAGAGTGT
Db      1002 TCTTGAACATGAACCTTCTCATGTGCTTTCATGACTGAGTCAATATATGAGGTG
QY      1058 GCTTCCTGCTGCTATGAAGAGCTCTGTGATCTTCAAGTGGCTCCAAAGTCCCTTG
Db      1062 GTTGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
QY      1118 CCTTCATCTGATCAAAAACATCTGACATCCGCTCAGTATTTGACCTTG-----
Db      1122 ACGTCTGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY      1172 ---TGGCTAATGAATCTTCCATGACAGCACTTCTGACTGAATTCGAGACAGTCTTC
Db      1182 GGAATCTTAATGAATATACAGAACTGGGGAGATCTGTGAABAAACAGAGACGGTACA
QY      1229 TGAAGTGTGAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Db      1242 CTGAGGTGTGAGACATGATATTTGTCAATCTGTGCATTAAGGCTGATATGCTTACA
QY      1289 GAGATGTGTGACCTGTCTACTTCTATGAGTTTCGGCAACCGGCTCAGTGTGTGAG
Db      1302 GAGATGAGGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY      1349 ACAGAAAGCCGCTTTTGTCAAAAGCCGACCAAGCTGATGAATCCGCTTGTGTGTGTG
Db      1362 ATAAAGATCAAGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY      1409 GTGCTTCTGAAAGGGGACATTTGTATGTTCAGAGAGCCAGGAGAGAGAAATAC
Db      1422 TTTGTTTCAAGTTTCTCAATCAATTAAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY      1469 TGAACCGAAGATGATGAATATCTGTGAGTCTTGTGCTGCAACCGGAAATCTTAATGGA
Db      1482 TGAACCAATCATGATGAGTACTGTGGGGCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY      1529 ACAGACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
Db      1542 AGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT

```

Search completed: June 15, 2005, 02:39:45
 Job time : 5801.46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2005, 12:34:34 ; Search time 290.702 Seconds
(without alignments)
9827.730 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 1746

Sequence: 1 atgcacagcaggactactc.....tttcttctgtgctccttga 1746

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1746	100.0	1746	US-10-023-515-3	Sequence 3, Appl1
2	1746	100.0	2158	US-10-023-515-1	Sequence 1, Appl1
3	508.6	23.1	965	US-09-799-451-155	Sequence 155, App
4	406.6	23.3	2117	US-09-949-016-3799	Sequence 3799, Ap
5	406.6	23.3	2169	US-09-949-016-555	Sequence 555, App
6	406.6	23.3	2191	US-09-595-6828-27	Sequence 27, Appl
7	362.6	20.8	1717	US-09-595-6828-20	Sequence 20, Appl
8	357	20.4	1701	US-09-264-737-3	Sequence 3, Appl1
9	139.4	8.0	1725	US-09-810-8618-5	Sequence 5, Appl1
10	139.4	8.0	1845	US-07-732-962A-1	Sequence 1, Appl1
11	139.4	8.0	1845	PCT-US92-06106-1	Sequence 1, Appl1
12	139.4	8.0	2158	US-09-949-016-1192	Sequence 1192, Ap
13	139.4	8.0	2158	US-09-949-016-1193	Sequence 1193, Ap
14	139.4	8.0	2256	US-08-318-826A-5	Sequence 5, Appl1
15	139.4	8.0	2256	US-08-370-156-1	Sequence 1, Appl1
16	139.4	8.0	2256	US-08-814-095-1	Sequence 1, Appl1
17	139.4	8.0	3016	US-08-318-826A-7	Sequence 7, Appl1
18	139.4	8.0	3016	US-08-370-156-5	Sequence 5, Appl1
19	139.4	8.0	3016	US-08-814-095-5	Sequence 5, Appl1
20	139.4	8.0	3096	US-08-318-826A-6	Sequence 6, Appl1
21	139.4	8.0	3096	US-08-370-156-3	Sequence 3, Appl1
22	139.4	8.0	3096	US-08-814-095-3	Sequence 3, Appl1
23	139.4	8.0	5767	US-09-810-861B-3	Sequence 3, Appl1
24	139.4	8.0	9885	US-09-949-016-12934	Sequence 12934, A
25	139.4	8.0	9885	US-09-949-016-12935	Sequence 12935, A
26	139.4	8.0	14446	US-09-810-861B-4	Sequence 4, Appl1
27	139.4	8.0	35060	US-08-614-095-7	Sequence 7, Appl1

28	128.6	7.4	12827	4	US-09-949-016-155A1	Sequence 155A1, A
29	123	7.0	1764	4	US-09-491-356C-3	Sequence 3, Appl1
30	123	7.0	2802	4	US-09-949-016-2515	Sequence 2515, Ap
31	123	7.0	2862	4	US-09-949-016-2516	Sequence 2516, Ap
32	123	7.0	4436	4	US-09-491-356C-6	Sequence 6, Appl1
33	121.4	7.0	4233	4	US-09-491-356C-5	Sequence 5, Appl1
34	105.6	6.0	836	4	US-09-270-767-16143	Sequence 861, Appl
35	105.6	6.0	836	4	US-09-270-767-16143	Sequence 16143, A
36	101.2	5.8	1905	3	US-09-347-878-33	Sequence 33, Appl
37	101.2	5.8	1907	1	US-08-462-884A-2	Sequence 2, Appl1
38	101.2	5.8	1908	1	US-08-461-881B-2	Sequence 2, Appl1
39	101.2	5.8	1908	2	US-09-123-960-2	Sequence 2, Appl1
40	101.2	5.8	2184	1	US-08-445-050-8	Sequence 8, Appl1
41	101.2	5.8	2184	4	US-08-204-691-8	Sequence 8, Appl1
42	101.2	5.8	2375	4	US-09-949-016-3976	Sequence 3976, Ap
43	101.2	5.8	2428	1	US-08-445-050-1	Sequence 1, Appl1
44	101.2	5.8	2428	1	US-08-204-691-1	Sequence 1, Appl1
45	101.2	5.8	2428	4	US-09-355-295B-2	Sequence 2, Appl1

ALIGNMENTS

```

RESULT 1
; Sequence 3, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Siles-Santiago, Inmaculada
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023, 515
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-023-515-3

Query Match          100.0%; Score 1746; DB 4; Length 1746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGCACAGGAGACTTACTTCATCTGCTTCACAGTGTGCTTTTCTGATTCACAGCC 60
DB      1 ATGCACAGGAGACTTACTTCATCTGCTTCACAGTGTGCTTTTCTGATTCACAGCC 60

QY      61 CTTTGGGACACAGACAGTGGGAAACTGGGCTTCTGCGAAGGGCACAGAGAAC 120
DB      61 CTTTGGGACACAGACAGTGGGAAACTGGGCTTCTGCGAAGGGCACAGAGAAC 120

QY      121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGACCTGTGCTGTG 180
DB      121 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGACCTGTGCTGTG 180

QY      122 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGACCTGTGCTGTG 180
DB      122 ACCAGGCTGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGAGACCTGTGCTGTG 180

QY      181 AAGTGTCTCTGAGATCCCTTCTGCTGCTCCCGCTGGATCCCTGCGATTACGAAC 240
DB      181 AAGTGTCTCTGAGATCCCTTCTGCTGCTCCCGCTGGATCCCTGCGATTACGAAC 240

QY      181 AAGTGTCTCTGAGATCCCTTCTGCTGCTCCCGCTGGATCCCTGCGATTACGAAC 240
DB      181 AAGTGTCTCTGAGATCCCTTCTGCTGCTCCCGCTGGATCCCTGCGATTACGAAC 240

QY      241 CCGGAGCTGGATGCGCCCTGGGATACTTGGAGAAAGCACTCTTAACCTAATTGTGC 300
DB      241 CCGGAGCTGGATGCGCCCTGGGATACTTGGAGAAAGCACTCTTAACCTAATTGTGC 300

QY      301 CTCGAGAACTGAGATGCTGCTTGAATCAACACATGCTCAAGGTGATACCGGAA 360
DB      301 CTCGAGAACTGAGATGCTGCTTGAATCAACACATGCTCAAGGTGATACCGGAA 360

```

Db 301 CTCGAGAACTCAGATGGCTGCTCTTAGATCAACAACATGCTCAAGTGCATTAACCCAAA 360
Qy 361 TTGGAGATGTCAGAAAGCTGGCTCTAAGTAACTTAATGGGCTGCCACGCCGATAGA 420
Db 361 TTGGAGATGTCAGAAAGCTGGCTCTAAGTAACTTAATGGGCTGCCACGCCGATAGA 420
Qy 421 GGCTCAAGCTCCCGCTTGGTGGTCCAGAGAGTGGCTTCAAGACTGGCTCAGCC 480
Db 421 GGCTCAAGCTCCCGCTTGGTGGTCCAGAGAGTGGCTTCAAGACTGGCTCAGCC 480
Qy 481 TCCATCTTTATAGAGTCCCGCTGGCTGGCTATGAGGACGCTGGTGGTGGTCCAG 540
Db 481 TCCATCTTTATAGAGTCCCGCTGGCTGGCTATGAGGACGCTGGTGGTGGTCCAG 540
Qy 541 TACCGGTAGAAATATTGGTTCTTCAACAATGGATGATGATGATGATGATGATGATG 600
Db 541 TACCGGTAGAAATATTGGTTCTTCAACAATGGATGATGATGATGATGATGATGATG 600
Qy 601 GCTTTCAAGGACAGAGTGGCTGCTGCTCTGGGTCCAGAGAAACATCGAGTCTTGGT 660
Db 601 GCTTTCAAGGACAGAGTGGCTGCTGCTCTGGGTCCAGAGAAACATCGAGTCTTGGT 660
Qy 661 GGGGACCCGAGCTCTGAGCAATCTTTGGGAGTCCCGGGAGCCATTAAGTCTTAACT 720
Db 661 GGGGACCCGAGCTCTGAGCAATCTTTGGGAGTCCCGGGAGCCATTAAGTCTTAACT 720
Qy 721 CTATACTGCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 780
Db 721 CTATACTGCTCCATGGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTG 780
Qy 781 GGCATATCCCTTACCTGAGGCTCATGATTAATGAGAAAGTGAAGACCTGACAGTGGT 840
Db 781 GGCATATCCCTTACCTGAGGCTCATGATTAATGAGAAAGTGAAGACCTGACAGTGGT 840
Qy 841 GCACATTTCTGCTGTAACATGCTGAGCTGAGAGCTGAGGCTGAGGCTGAGGAGCA 900
Db 841 GCACATTTCTGCTGTAACATGCTGAGCTGAGAGCTGAGGCTGAGGCTGAGGAGCA 900
Qy 901 AAACCTCCAGAGAGTGGCTGAGCCCTGAGCCAGCAAAACAAAGTCTTCACTGAGTGGT 960
Db 901 AAACCTCCAGAGAGTGGCTGAGCCCTGAGCCAGCAAAACAAAGTCTTCACTGAGTGGT 960
Qy 961 GATGCTGCTTCTTCTCTAATGAGGCTCTAGATCTAATGCTCAGAAAGCATTTAAACA 1020
Db 961 GATGCTGCTTCTTCTCTAATGAGGCTCTAGATCTAATGCTCAGAAAGCATTTAAACA 1020
Qy 1021 ATTCTTCATCATGAGATCAATTAACACAGAGTGGCTTCTGCTCATGAGAGAG 1080
Db 1021 ATTCTTCATCATGAGATCAATTAACACAGAGTGGCTTCTGCTCATGAGAGAG 1080
Qy 1081 GCTCTGAGATCTCAGAGTCCCAAAAGTCCCTTGGCTCCATCTGATCAAAACATC 1140
Db 1081 GCTCTGAGATCTCAGAGTCCCAAAAGTCCCTTGGCTCCATCTGATCAAAACATC 1140
Qy 1141 CTGACATATCCCGCTCAGATATTTGCACTTGGCTAATGAATCTTCCATGAGACAC 1200
Db 1141 CTGACATATCCCGCTCAGATATTTGCACTTGGCTAATGAATCTTCCATGAGACAC 1200
Qy 1201 TCCCTGACTGAATCCAGACAGTCTTGGAGCTTGGAGTGTCTTCTTGTGCTC 1260
Db 1201 TCCCTGACTGAATCCAGACAGTCTTGGAGCTTGGAGTGTCTTCTTGTGCTC 1260
Qy 1261 CTTGCACTGATCAAGCTCGATATCAAGAGATGCTGATGATGATGATGATGATGATG 1320
Db 1261 CTTGCACTGATCAAGCTCGATATCAAGAGATGCTGATGATGATGATGATGATGATG 1320
Qy 1321 TTTTGGGACCGGCTCAGTGTGTTGAAGACAGAGCCGCTTTTGTCAAGCCGACAC 1380
Db 1321 TTTTGGGACCGGCTCAGTGTGTTGAAGACAGAGCCGCTTTTGTCAAGCCGACAC 1380
Qy 1381 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
Db 1381 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440

Qy 1441 GAAGAGCCACGAGAGGAGAGATTTAAGTACCGGAGAGATGATAATTAAGGCTACC 1500
Db 1441 GAAGAGCCACGAGAGGAGAGATTTAAGTACCGGAGAGATGATAATTAAGGCTACC 1500
Qy 1501 TTTTGGTCCAGCCGAGATCTTAATGAGAACGACTGTCTGTGGCAAGCTTAATATG 1560
Db 1501 TTTTGGTCCAGCCGAGATCTTAATGAGAACGACTGTCTGTGGCAAGCTTAATATG 1560
Qy 1561 ACTGAGCAGTACCTCAGCTGAGCTTTGAACATGAGCTCCGACAGAGACTAAAGAAC 1620
Db 1561 ACTGAGCAGTACCTCAGCTGAGCTTTGAACATGAGCTCCGACAGAGACTAAAGAAC 1620
Qy 1621 CGGATGATTTTGGACAGACACATCCCGTGAATCTGTGCTGCTCCGACATGCTCAC 1680
Db 1621 CGGATGATTTTGGACAGACACATCCCGTGAATCTGTGCTGCTCCGACATGCTCAC 1680
Qy 1681 AGTCTCTTCTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Db 1681 AGTCTCTTCTCTCTTAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740
Qy 1741 CCTTGA 1746
Db 1741 CCTTGA 1746

RESULT 2
US-10-023-515-1
; Sequence 1, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (96)...(1838)
US-10-023-515-1

Query Match 100.0%; Score 1746; DB 4; Length 2158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGSCACAGGAGCTTATCATCTGCTTCAATGATGATGATGATGATGATGATGATGATG 60
Db 96 ATGSCACAGGAGCTTATCATCTGCTTCAATGATGATGATGATGATGATGATGATGATG 155
Qy 61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGACACAGAGAAC 120
Db 156 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGACACAGAGAAC 215
Qy 121 ACCAGGCTGGGATGATTTCAAGGCAAGACATGCTGTGGGAAGCCCTGTGCTGTG 180
Db 216 ACCAGGCTGGGATGATTTCAAGGCAAGACATGCTGTGGGAAGCCCTGTGCTGTG 275
Qy 181 AACGTTCTCTGAGAGTCCCTTGTGCTGCTCCCGCTGGGATCCCTGGAATTAATGAAAC 240
Db 276 AACGTTCTCTGAGAGTCCCTTGTGCTGCTCCCGCTGGGATCCCTGGAATTAATGAAAC 335

241 CCGACGCTGACCGCCCTGGAGTAATTCGAGAGACCACTCCCTAATTTGTC 300
Db CCGAGCCTGATGCCCTGGGATACTTGGAGAGACCACTCTAATTTGTC 395
Qy 301 CTCGAAATCGAGTGGCTGCTTTAGTCAACATGCTCAAGGTGATTAACCGAAA 360
Db CTCGAAATCGAGTGGCTGCTTTAGTCAACATGCTCAAGGTGATTAACCGAAA 455
Qy 361 TTCCGAGGTGAGAGACTGCTCACTGAACTATTAAGGCGCCGACCGCATCA 420
Db TTCCGAGGTGAGAGACTGCTCACTGAACTATTAAGGCGCCGACCGCATCA 515
Qy 421 GGGCTCAAGCTCCCGCTTGGTGTGTTCCGAGAGGTGCTTCAAGATGGCTCAGCC 480
Db GGGCTCAAGCTCCCGCTTGGTGTGTTCCGAGAGGTGCTTCAAGATGGCTCAGCC 575
Qy 481 TCCATCTTGAATGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db TCCATCTTGAATGGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635
Qy 541 TACCGGCTAGGAATTTGGTCTTTCACCAATGGGATACAGATGCTCCGGGAACTGG 600
Db TACCGGCTAGGAATTTGGTCTTTCACCAATGGGATACAGATGCTCCGGGAACTGG 695
Qy 601 GCTTCAAGACCAAGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db GCTTCAAGACCAAGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
Qy 661 GGGGACCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db GGGGACCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
Qy 721 CTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db CTATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
Qy 781 GGCATCATCCCTTACCTGAGAGGCCATGATTAAGAGAGAGAGAGAGAGAGAGAG 840
Db GGCATCATCCCTTACCTGAGAGGCCATGATTAAGAGAGAGAGAGAGAGAGAGAG 935
Qy 841 GCACATTTCTGTGTAAACATGCTGCTGAGAGGCCCTGCTGAGGTGCTGAGAGCA 900
Db GCACATTTCTGTGTAAACATGCTGCTGAGAGGCCCTGCTGAGGTGCTGAGAGCA 995
Qy 901 AAACCTTCAAGAGAGCTGCTGAGCCCTGAGCCGAGAGAGAGAGAGAGAGAGAG 960
Db AAACCTTCAAGAGAGCTGCTGAGCCCTGAGCCGAGAGAGAGAGAGAGAGAGAG 1055
Qy 961 GATGATGCTTCTTCTTAATGAGGCTGATGATGATGATGATGATGATGATGATGAT 1020
Db GATGATGCTTCTTCTTAATGAGGCTGATGATGATGATGATGATGATGATGATGAT 1115
Qy 1021 ATTCTTCAATCATGAGATCAATAACCAAGAGTGGCTTCTGCTGCTGATGAGAGAG 1080
Db ATTCTTCAATCATGAGATCAATAACCAAGAGTGGCTTCTGCTGCTGATGAGAGAG 1175
Qy 1116 ATTCTTCAATCATGAGATCAATAACCAAGAGTGGCTTCTGCTGCTGATGAGAGAG 1175
Db ATTCTTCAATCATGAGATCAATAACCAAGAGTGGCTTCTGCTGCTGATGAGAGAG 1235
Qy 1081 GCTCTGAGATCTGAGTGGCTTCAAGAGTGGCTTCTGCTGCTGATGAGAGAGAG 1140
Db GCTCTGAGATCTGAGTGGCTTCAAGAGTGGCTTCTGCTGCTGATGAGAGAGAG 1235
Qy 1176 GCTCTGAGATCTGAGTGGCTTCAAGAGTGGCTTCTGCTGCTGATGAGAGAGAG 1235
Db GCTCTGAGATCTGAGTGGCTTCAAGAGTGGCTTCTGCTGCTGATGAGAGAGAG 1295
Qy 1141 CTGACATCCCGGCTCAGATATTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db CTGACATCCCGGCTCAGATATTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1295
Qy 1201 TCCCTGATCAATCGAGATCGAGATCTTGAAGTGGCTTCTGCTGCTGCTGCTGCT 1260
Db TCCCTGATCAATCGAGATCGAGATCTTGAAGTGGCTTCTGCTGCTGCTGCTGCT 1355
Qy 1261 CCTGACATGATCAAGCTGCTGATTAACAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db CCTGACATGATCAAGCTGCTGATTAACAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 1415
Qy 1356 CCTGACATGATCAAGCTGCTGATTAACAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 1415
Db CCTGACATGATCAAGCTGCTGATTAACAGAGATGCTGCTGCTGCTGCTGCTGCTGCT 1415
Qy 1321 TTTCGGACACGGGCTCAAGTCTTGAAGACGAGAGCCGGCTTTTGTCAAGCCGACAC 1380

Db 1416 TTTCGGACACGGGCTCAAGTCTTGAAGACGAGAGCCGGCTTTTGTCAAGCCGACAC 1475
Qy 1381 GCTGATGAAGTCCCTTTGTGTTCGGTGTGCTTCTGTAAGGGGAGACATTTATGCTC 1440
Db 1476 GCTGATGAAGTCCCTTTGTGTTCGGTGTGCTTCTGTAAGGGGAGACATTTATGCTC 1535
Qy 1441 GAAGAGCCACGAG 1500
Db GAAGAGCCACGAG 1595
Qy 1501 TTTCCTGAAACCGGAGATCTTAATGAGAACGACCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db TTTCCTGAAACCGGAGATCTTAATGAGAACGACCTGCTGCTGCTGCTGCTGCTGCTGCT 1655
Qy 1561 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCTTCCGACAGAGATCAAGAGACCG 1620
Db ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCTTCCGACAGAGATCAAGAGACCG 1715
Qy 1621 CGGGTGAATTTTGAACGACACATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db CGGGTGAATTTTGAACGACACATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1775
Qy 1681 AGTCTCTTCTTCTTAACTTCT 1740
Db AGTCTCTTCTTCTTAACTTCT 1835
Qy 1741 CTTTGA 1746
Db 1836 CTTTGA 1841

RESULT 3
US-09-799-451-155
; Sequence 155, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yuning
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunhui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pl_FL_genes Version 2.0
; SEQ ID NO 155
; LENGTH: 965
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(925)
; US-09-799-451-155

Query Match 29.1%; Score 508.6; DB 4; Length 965;
Best Local Similarity 81.0%; Pred. No. 3.2e-150;

Matches 675; Conservative 0; Mismatches 4; Indels 154; Gaps 1;

QY 91 GGGGCTTCTGTGAAGGCGCCAGAGAAACCCAGGCTGGATGATTCAGGGCAACAA 150
 Db 287 GGGGCTTCTGTGAAGGCGCCAGAGAAACCCAGGCTGGATGATTCAGGGCAACAA 150
 QY 151 GTCACTGTGTGGGAAGCCCTGTGCTCTGTGAACGTTTCTCGAGTCCCTTGTGCT 210
 Db 347 GTCACTGTGTGGGAAGCCCTGTGCTCTGTGAACGTTTCTCGAGTCCCTTGTGCT 210
 QY 211 CCCCCGTGGGAATCCCTGTGCTCTGTGAACGTTTCTCGAGTCCCTTGTGCT 270
 Db 407 CCCCCGTGGGAATCCCTGTGCTCTGTGAACGTTTCTCGAGTCCCTTGTGCT 270
 QY 271 CGAAGAGCCACTCTTACCTTAATTTGTGCTCTGAAGCCGAGCTGCTGCTTAACT 330
 Db 467 CGAAGAGCCACTCTTACCTTAATTTGTGCTCTGAAGCCGAGCTGCTGCTTAACT 330
 QY 331 CAACATGCTCAAGGTGCTTAACCCGAAATTCGAGTGTCAAGAACTGCTTACCTG 390
 Db 527 CAACATGCTCAAGGTGCTTAACCCGAAATTCGAGTGTCAAGAACTGCTTACCTG 390
 QY 391 AACATTAAGCGCTGCTCCAGCCGATTAAGGCTTCAAGTCCCTGCTTGTGCTTTC 450
 Db 587 AACATTAAGCGCTGCTCCAGCCGATTAAGGCTTCAAGTCCCTGCTTGTGCTTTC 450
 QY 451 CCAGAGAGTGTCTTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
 Db 647 CCAGAGAGTGTCTTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510
 QY 511 TATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
 Db 707 TATGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
 QY 571 ACATGGAGTCAAGATGCTCCGGGAACCTGGGCTTCAAGAGCAAGTGTGCTGCTGCT 630
 Db 766 ----- 765
 QY 631 TGGGTCCAGAGAGATCGAGTTCTTGGTGGGAGCCCAAGCTGTGACCATTTTGGC 690
 Db 766 ----- 765
 QY 691 GAGTCCGGGGAGCCATAAGTCTTCTAGTCTTAACTGCTGCTCCATGCGCAAGGCTTAA 750
 Db 766 ----- -ATACTGCTCCCATGCGCAAGGCTTAA 792
 QY 751 TTCCCAAGAGCATATGAGAGTGGGAGGAGCATCCCTTACCTGAGAGCCCATGAT 810
 Db 793 TTCCCAAGAGCATATGAGAGTGGGAGGAGCATCCCTTACCTGAGAGCCCATGAT 852
 QY 811 TATGAGAGTGAAGAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870
 Db 853 TATGAGAGTGAAGAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
 QY 871 TCTGAGGCGCTGCTGAGGCTGCTGAGCAAAACCTTCCAGAGGCTGCTGAC 923
 Db 913 TCTGAGGCGCTGCTGAGGCTGCTGAGCAAAACCTTCCAGAGGCTGCTGAC 965

RESULT 4
 US-09-949-016-3799
 ; Sequence 3799, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: CU001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 3799
 ; LENGTH: 2117
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-3799

Query Match 23.3%; Score 406.6; DB 4; Length 2117;
 Best Local Similarity 56.5%; Pred. No. 1.2e-117;
 Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;

QY 116 GGAAACAGGCTGGATGATTCAGGCAAGCAAGTCACTGTGGAAGCCCTGCTGCT 175
 Db 157 GGAAACAGGCTGGATGATTCAGGCAAGCAAGTCACTGTGGAAGCCCTGCTGCTGCT 216
 QY 176 CTGTGAAGTGTCTCTGGAAGTCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235
 Db 217 GGGTCCAAACCTCTGGGAATTCATTTGCCAAGCCACTTAAAGTCCGCTGGAATTG 276
 QY 236 GGAACCCGAGCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
 Db 277 CACCCCTGAGCCCTCTGAATCTTGAAGTGTGAGGAGTGAAGCAACCATTCGCTCA 336
 QY 296 TGTGCTTCAAGAGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 352
 Db 337 TGTGCTTCAAGAGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 396
 QY 353 ACCGGAATTCGAGTGTCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412
 Db 397 TCCCTTCCGACTCATGTGTGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 456
 QY 413 CCGATACAGGCTCCCAAGCTCCCGCTTGGTGTGCTTCCAGAGGAGGCTTCAAGAGCT 472
 Db 457 GCGATGAAGGCTCTMAACTCCCGGTATGTGTGATTCACAGGTGGTCCCTGTTTGTG 516
 QY 473 GCTAGGCTTCACTTTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532
 Db 517 GATGAGCTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
 QY 533 TCGTCCAGTACCGGCTAAGAAATTTGTGCTTCTTCAACATGAGATACAGTCTGCTGCT 592
 Db 577 TCAATCAGTACCGGCTAAGAAATTTGTGCTTCTTCAACATGAGATACAGTCTGCTGCT 636
 QY 593 GGAACCTGAGCTTCAAGAGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 652
 Db 637 GGAACCTGAGCTTCAAGAGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
 QY 653 TCTTGGTGGGAGCCCAAGCTGTGTACCAATCTTGGGCAAGTCCGAGGAGCCATTAAGT 712
 Db 697 ACTTTGAGGCAACCTTGAACCGGTGTCAATTTTGGGCAAGTCCGAGGAGCCATTAAGT 756
 QY 713 TTTCTAGTCTTATATGCTGCTCCAGAGGCTTATTCACAAAGCATCATGAGAA 772
 Db 757 TGTCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
 QY 773 GTGGGAGGCAATCTTCACTGAGAGCCCATGATTAAGAAAGATGAGAGCTGCTGCTGCT 832
 Db 817 GTGGGAGGCAATCTTCACTGAGAGCCCATGATTAAGAAAGATGAGAGCTGCTGCTGCT 870
 QY 833 AGGTGAGGCAATCTTCACTGAGAGCCCATGATTAAGAAAGATGAGAGCTGCTGCTGCT 892
 Db 871 CGGTGAGGCAATCTTCACTGAGAGCCCATGATTAAGAAAGATGAGAGCTGCTGCTGCT 930
 QY 893 TGAGGCAAAACCTTCAAGAGAGTGTGAGCCCTGAGGAGAAACAAAGTCTTCACTGCT 952
 Db 931 TGAGGAGCAAGAGTGAAGAGAGATTTTGAATTAAGAGCTTCAAGATGATTCCTGCT 990
 QY 953 GAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012

Db 991 GAGTGTGATGAGGGCTTCTCTGCCAGGCAACCCAGAGCTGTGCGCTTCCGCACT 1050
Qy 1013 TTAAGCAATTCCTTCATCATCGAGTCAATTAACGAGTGTGCTTCGCTGCCA 1072
Db 1051 TTCAGCTGTCTCCCTGAGCATTTGGTGTCAACAAGTAATTCGCTGGCTCATCCCA 1110
Qy 1073 TGA-----AGGAGCTCCTGAGATCTCGAGTGTCCAAAGATCCCTTGGCCC 1120
Db 1111 AGGTATAGAGATCTATATACCCAGAGAAATGAGACAGAGGCTCCAGGCTGTCTC 1170
Qy 1121 TCCATCTGATACAAAACATCTGCAACATCCGCTCAGTATTGCACTTTGGGCTAATG 1180
Db 1171 TGCAGAAATGTTAAGCTGTGATGTGTCTTCAATTGTGTGACCTGTGAGGGAGG 1230
Qy 1181 AATATCTTCATGACAGAGCACTCCCTGACCTGAATTCGAGACAGTCTTCTGACCTTGG 1240
Db 1231 AGTACATTTGGGAGCAATGGGGATCCCCAGACCCCTCCAGGCGAGTTCCAGAGAAATGATG 1290
Qy 1241 GAGATGTCTTTTGTGCTCCTGCACTGATCAAGCTCGATATGACAGAGATGCTGTG 1300
Db 1291 CGGACTCGATGTTTGTGATCCCTGCACTCCAGTACATTTTCACTG---TTCGGGG 1347
Qy 1301 CACTGTCTACTTCTATATGATTTGGGCAACGGGCTCAGTCTTTGAAACAGAAAGCCGG 1360
Db 1348 CCCCTGTGTACTTCTACGAGTTCCAGCATCAGCCAGCTGGCTCAAGAAATCAGGGCAC 1407
Qy 1361 CTTTGTCAAGCCGCAACGCTGATGAAATCCGCTTGTGTGCTGTGTGCTTCTTGA 1420
Db 1408 CGCATATGAGGCAACATGATGATGAGCTTCTTTGTTTCAAGATTTCTTTGGGG 1467
Qy 1421 AGGGGCAATTTGTTATGTTTCAAGAGCCAGGAGAGAGAAATTACTGAGCCGGAAGA 1480
Db 1468 GCACTACATTTAAAT-----CACTAGAGAAAGAGAGACAGCTTAGCAGAGAGA 1515
Qy 1481 TGATGAAATCTGGGCTACCTTTGCTCGAACCGGGAATCTTAATGGAACGACCTGTCTC 1540
Db 1516 TGATGAAATCTGGGCTACCTTTGCTCGAACCGGGAATCTTAATGGAACGACCTGTCTC 1575
Qy 1541 TGTGGCCAGCTTATATATGATGACGAGTACCTTCACTGATGATCTTGAACATGAGACTCG 1600
Db 1576 ACTGCGCTGTGTGACCAAGAGAGAGCAATCTTCACTGATGATCTTGAACATGAGACTCG 1635
Qy 1601 GACAGAGACTCAAGAACCGCGGTGATTTTGGACCGACCATCCCCC 1651
Db 1636 GCCGGGCTGTGAGGCCCAAGGCTCCAGTTCTGGAAGAGGCGCTGCCCC 1686

RESULT 5
US-09-949-016-555
; Sequence 555, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 555
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-555

Query Match 23.3%; Score 406.6; DB 4; Length 2169;

Best Local Similarity 56.5%; Pred. No. 1.2e-117;
Matches 876; Conservative 0; Mismatches 639; Indels 36; Gaps 5;
Qy 116 GGAACACCAAGGCTGGAGATGATTCAGGCGCAAGCAAGTCACTGTGCGGAACCCCTGTGC 175
Db 157 GGACCAACAACAGGGGAGGCTGGGGAGCTTGTCAATGTGAAGGGCGCAATATGCCG 216
Qy 176 CTGTGAACGTGTTCTCGGAGTCCCTTGTGCTCTCCCGCTGGAGATCCCTGCATTTA 235
Db 217 GGGTCCAAACCTTCTGGGAATTCATTTGCCAACCACTCTAGTCCGCTGCCATTTTG 276
Qy 236 CGAACCCGACGCTGCATGCGCTGGGATTACTTCGAGAACCACTTCTTACCTTAAT 295
Db 277 CACCCCTGAGGCCCTGAAATCTTGAATGTGTGAGGGATGAACCAACCATCCGCCCA 336
Qy 286 TGTGCTCCAGAACTCAAGTGGCTGTCTTGA---TCAACAATCTCAAGTGCAT 352
Db 337 TGTGTCTACAGGACCTCAACCGAGTGAAGTCAAGATTTCTTAAGCAATTCATGACCT 396
Qy 353 ACCGAAATTCGAGAGTCAAGAGACTGCTCTTACCTGAACATCTATGCGCTGCCACG 412
Db 397 TCCCTTCCGACTCATGTCTGAGAGCTGCTGTACTTCAAGATCTTACAGCGCGGCATA 456
Qy 413 CCGATACAGGCTCCAAAGCTCCCGCTTGTGTGTGTGTTCCAGAGATGTCTTCAAGACTG 472
Db 457 GCATGAAGGCTCTAACCTGCGGTGATGTGTGATCAGCGTGTGTGGCTTGTGTTTG 516
Qy 473 GCTCAGCTCCATCTTATATGAGTTCGCGCTGTGCTGCTTATAGAGACGTGTGTGTG 532
Db 517 GCATGCTTCTTGTATATGATGTTCATGTGCTGCTGTGAGAACTGTGTGTGTCA 576
Qy 533 TGTGTCAGTACCGGTAGAAATTTGTTTCTTCAACCATGAGATGACATGTCTCGG 592
Db 577 TCATCAATGACCGCTGTGTGTCTGTGCTTCTTGAAGCTGGAACAGACGCAACCG 636
Qy 593 GGAATGCGCTTCAAGAGCAAGTGTGCTGTCTCTGTGGTTCAGAGAAATCTGAGT 652
Db 637 GCAATGGGGCTACCTGAGCAAGGTGTGCTACAGCTGGGTCCAGAGAAATATGCC 696
Qy 653 TTTTGGTGGGAGCCCAAGCTGTGACCATTTTGGGAGTCCCGGGAGCATTAATG 712
Db 697 ACTTGGAGGCAACCTGACCGGTGACCATTTTGGGAGTCTCGGGTGGCAGAGTG 756
Qy 713 TTTTCACTTATACGTCTCCATGAGCCAAAGGCTTATTCAGAAAGCATATGAGA 772
Db 757 TGTCTTCCGTTGTGTGTCCTCCATATCCCAAGACTCTTCCAGAGCATATGAGA 816
Qy 773 GTGGGTGGCCATCATCCCTTACCTGAGGCCCATGATTAAGAAAGTGAAGACCTGC 832
Db 817 GTGGGTGGCCCTCTGCGCGGCTCATTTGCCAGCTGACGTGA-----TGCATCTCA 870
Qy 833 AGGTGTTGACATTTCTGTGTAACAATGCTCAGACTTGAAGCCCTGTGAGAGTGC 892
Db 871 CGGTGTGGCAACCTGTCTGCTGTGACAAATTTGACCTTGAAGCCCTGTGTGGCTGC 930
Qy 893 TGAGGCAAAACCTCCAGAGGCTGTGACCTCAGGCAAGAAACAAAGTCTTCACTC 952
Db 931 TCGGGGCAAGATTAAGAGAGATTTCTGCAATTAACAGCTTTCAAGATATCCCG 990
Qy 953 GAGTGTGATGAGTCTTCTTCTTAATGAGCTCTGATCTATGTCTCAGAAAGCAT 1012
Db 991 GAGTGTGATGAGGCTCTTCTGCGCCAGGACCCCAAGAGTGTGGCTGTGCGACT 1050
Qy 1013 TTAAGCAATTCCTTCATCATCGAGTCAATTAACGAGTGTGCTTCTGTGCTCTA 1072
Db 1051 TTCAGCTGTCTCCCTGAGCATTTGGTGTCAACAAGTAATTCGCTGGCTCATCCCA 1110
Qy 1073 TGA-----AGGAGCTCCTGAGATCTCGAGTGTCCAAAGATCCCTTGGCCC 1120
Db 1111 AGGTATAGAGATCTATATACCCAGAGAAATGAGACAGAGGCTCCAGGCTGTCTC 1170
Qy 1121 TCCATCTGATACAAAACATCTGCAACATCCGCTCAGTATTGCACTTTGGGCTAATG 1180

Db 1370 CCCCTGTGTACTTCTACGAGTTCCAGATCAACGCCAGCTGCTCAAGAACTACAGGCCAC 1429
Qy 1361 CTTTGTCAAGCCGACCAACGCTGATGAAGTCGCTTTGTGTGGTGGTGGCTTCTCTGA 1420
Db 1430 GCGACATGAAGGACAGACATGATGATGAGCTTCTTTGTTCAGAACTTTCTTTGGGG 1489
Qy 1421 AGGGGACATTTGTATGTTTCGAAGAGCCAGAGAGAGAGAACTTACTGAGCCGAGAGA 1480
Db 1490 GCAACTACATTAATTT-----CACTGAGGAGAGAGAGAGAGCTAAGCAGAGAGA 1537
Qy 1481 TGATGAATATCTGGGCTACCTTTGCTCCAAACCGGAACTCTAATGAGAAAGACCTGTCTC 1540
Db 1538 TGATGAATATCTGGGCTACCTTTGCTCCAAACCGGAACTCTAATGAGAAAGACCTGTCTC 1597
Qy 1541 TGTGGCCAGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1600
Db 1598 ACTGCGCCCTGTTGCAACGAG 1657
Qy 1601 GACAGAGACTCAAGAAACCGCGGTGATTTTTCGACAGACCAATCCCTC 1651
Db 1658 GCGCGGCTCTGAAGGCCACAGAGCTCAAGTTCTGGAAGAGGCGCTGCCCC 1708

RESULT 7
US-09-595-682B-20
Sequence 20, Application US/09595682B
Patent No. 6800483
GENERAL INFORMATION:
APPLICANT: Danks, Mary K.
APPLICANT: Porter, Philip M.
APPLICANT: Houghton, Peter J.
TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
TITLE OF INVENTION: Tumor Cells
FILE REFERENCE: SU-0005
CURRENT APPLICATION NUMBER: US/09/595,682B
CURRENT FILING DATE: 2000-01-16
PRIOR APPLICATION NUMBER: 60/075,258
PRIOR FILING DATE: 1998-02-19
PRIOR APPLICATION NUMBER: PCT/US99/03171
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 1717
TYPE: DNA
ORGANISM: *Oryctolagus cuniculus*
US-09-595-682B-20

Query Match 20.8% Score 362.6; DB 4; Length 1717;
Best Local Similarity 54.8%; Pred. No. 8.8e-104;
Matches 880; Conservative 0; Mismatches 649; Indels 78; Gaps 5;
Qy 91 GGGCTTGTGTCAGAGGCGCAACAGAGAAACCAAGCTGGATGATTCAGGGCAAGCA 150
Db 62 GGGGACCGGTGCAACACGTGTGTAAGTACGTGCAATGGCAAGTCTGGGAAAGTTC 121
Qy 151 GTCACTGTGCTGGGAAGCCCTGTGCTGAACTGTTCTCTGAGTCCCTTTGCTGCT 210
Db 122 GTCACTGTGCTGGGAAGCCCTGTGCTGAACTGTTCTCTGAGTCCCTTTGCTGCT 181
Qy 211 CCCCCGTGGGATCCCTGGAATTTACAAACCGGAGCTGATGCGCCCTGGGATTAATTTG 270
Db 182 CCCCCGTGGGATCCCTGGAATTTACAAACCGGAGCTGATGCGCCCTGGGATTAATTTG 241
Qy 271 CGAAGACCACTCTTACCTTAATTTGTGCTTCAAGAACTCTCAAGTGGCTGCTC 324
Db 242 AAGAACCACTCTTACCTTCCATGCTGCTCCAGAGAGAGATTCAGAGGATGATGCTC 301
Qy 325 TTAGTCAACACATGCTCAAGTGCATTAACCGAAATTCGAGTGTCAAGAACTGCTC 384
Db 302 TCGAGCTCTTCAACCAAGAAAGAAACATCCCTTAACTTTCTGAAAGACTGCTCCT 361
Qy 385 TACCTGAACATCTATGCGCTGCCACGCGGATACAGGCTCAAGCTCCCGTCTTGGTG 444

Db 362 TACCTGAATATTTACACCCCTGCTGACCTGACAAAGAGAGGAGCTGCGCTGATGCTG 421
Qy 445 TGGTCCAGAGAGGCTCTCAAGCTGCTCAGCTCCTCATCTTTGATGGGTCGCGCCG 504
Db 422 TGGATCTCAAGAGAGGCTCTGATGCTGAGAGATCACTTAAGATGGCTGCTGCTT 481
Qy 505 GCTGCTTGAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
Db 482 TCTGCCCATGAGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
Qy 555 TTACCAACATGAGATCAGATGCTCCTGAGAACTGAGGCTTCAAGAGCAAGTGGCTGCT 624
Db 542 TTACCAACATGAGATCAGATGCTCCTGAGAACTGAGGCTTCAAGAGCAAGTGGCTGCT 601
Qy 625 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
Db 602 CTGCGGTGGCTCAGAGCAACATTCCTCAACTTTGAGAGGAGCCAGGCTCTGTGACCATC 661
Qy 685 TTTGCGAGTCCGCGGAGCCATAAGTGTCTAGCTCTTAATCTGCTTCCATGAGCCAA 744
Db 662 TTTGAGAGTCAAGAGAGGCTCAAGTGTCTTATCTTATCTTATCTTATCTTATCTTATCT 721
Qy 745 GGTATATTCACAAAGCCATCATGAGAGTGGGAGTGGCCATCATCTTACTGAGGCC 804
Db 722 AATCTCTTCATCAGCAATTTCCGAGAGTGGCGTGGCTCTTCCAGTCTTCTTCAAG 781
Qy 805 CATGATTTAG 864
Db 782 AAGAACCAACAGTCTTGGCTGAGAAATTCGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 841
Qy 865 TCAGACTGAGAGGCTGCTGAGTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923
Db 842 TCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901
Qy 924 -----CTTCAGCAGAGAAACAAAGTCT 945
Db 902 TTGAATATGAATTTATGCTCTAGATCTAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 961
Qy 946 TTCACTCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005
Db 962 CTGACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
Qy 1006 AAGCAATTTAAAGCAATCTTCCATCATCGAGTCAATTAACAGAGAGTGGCTTCTG 1065
Db 1022 AAGAAATACACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
Qy 1066 CTGCTATGAAGAGAGCTCTGAGATCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1114
Db 1082 ATCCCAATGCAAAATCTGAGCTATCACTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
Qy 1115 -TTGCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1173
Db 1142 ACAGAACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201
Qy 1174 GCTAATG---AATATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1230
Db 1202 GCACTGAGAGATTTGAG 1261
Qy 1231 GACTTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1290
Db 1262 GAGATGCTTGAAGATTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1321
Qy 1291 GATGCTGAGCACTGCTACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1350
Db 1322 GATGCTGAG 1381
Qy 1351 ACGAAGCGGCTTTTGTCAAG 1410
Db 1382 ATGAGACCAAG 1441
Qy 1411 GCTTCTGAG 1470

Db 1442 CCGTTTTTAA-----AGAGGTCCACAGAAAGAGATCAAACTG 1483
QY 1471 AGCCGGAAGATGATGAATACTGGGCTACCTTGTCTGGAACGGGAATCTTAATGGAAAC 1530
Db 1484 AGGAAGATGGTGAATGAATACTGGGCTACCTTGTCTGGAATGGAAATCCCAATGAGAA 1543
QY 1531 GACCTGTCTGTGGCCAGCTTAACTGATGAGAGTACCTCCAGCTGACCTTGAAC 1590
Db 1544 GGGCTTCTCAATGGCCAGCATATGACTACAGGAAGTTACCTGAGATGGAGCAAC 1603
QY 1591 ATGAGCTGGAGAGACTCAAGAACCGCGGTGATTTTGGAC 1637
Db 1604 ACCCAGGACCCAGAAATGAAAGCAAGAAAGTGGCTTTTGAC 1650

RESULT 8
US-09-264-737-3
; Sequence 3, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ruff, Thomas G.
; TITLE OF INVENTION: Engineering of Plant Resistance to Pyridines via
; FILE REFERENCE: 38-21(10551) RLE3 Pyridine Tolerance
; CURRENT APPLICATION NUMBER: US/09/264,737A
; EARLIER FILING DATE: 1999-03-09
; EARLIER APPLICATION NUMBER: 60/077,377
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1701
; TYPE: DNA
; ORGANISM: Rabbit
US-09-264-737-3

Query Match 20.4%; Score 357; DB 3; Length 1701;
Best Local Similarity 54.8%; Pred. No. 5.2e-102;
Matches 866; Conservative 0; Mismatches 635; Indels 78; Gaps 5;

QY 119 ACACCAAGCTGGAGATGATTCAGGCAAGCAAGTCACTGTCTGGAGCCCTGTGCTG 178
Db 83 ACACGTGAAGGGAAGTCTCTGGGAAGTTCGTACGCTTGAAGATTTGCACAGCCG 142
QY 179 TGAAGTGTCTGTGAGTCTCTTGTCTGCTCCCGCTGGAGTCCCTGGATTTAACA 238
Db 143 TGGCGGTCTCTGGAGTCTCTTGTCTGCTCCCGCTGGAGTCCCTGGATTTAACA 202
QY 239 ACCCGAGCTGCAATCCGCTGGAGTAACTTGCAGAAAGCACTCTCTAATTTGT 298
Db 203 CACCAAGCTTGAATCATGAGCCAGTGAAGAACACCTCTTACCTCCCATGT 262
QY 299 GCTTCCAGAAC-----TCAGAGTGGCTGCTTGAATCAACATGCTCAAGGTGACT 352
Db 263 GCTCCCAAGAGAGATCATGAGCCAGTGAAGAACACCTCTTACCTCCCATGT 322
QY 353 ACCCGAAATTCGAGTGTCAAGAACTGCTCTAATCTGAACATCTATGCTGCGCAG 412
Db 323 ACATCCCTTAAATTTCGAAAGCTGCTTACCTGAATAATTTACACCTGCTGACC 382
QY 413 CCGATACAGGCTCCAGCTCCCGCTTGTGTGTGTTCCAGAGGTGCTTCAAGACTG 472
Db 383 TGAACAAGAGAGAGAGCTGCGGTGATGTGTGATTCATGAGGTGTGTGATGTG 442
QY 473 GCTCAGCTTCATTTTGTGTGTGCTGCGCTGTGCTTATGAGAGTGTGTTGTG 532
Db 443 GTGAGAGATCAACTATGATGCTGCTGCTTCTTCTGCAATGAAAGTGTGTGTGA 502
QY 533 TCGTCCAGTACCGGCTGAGAAATTTGTTTCTTCAACAATGAGATCAGATGCTCCG 592
Db 503 CCAATTCAGTACCGCTGGGATCTGGAGATTTTCAGCAAGAGATGAGCAAGCCGAG 562

QY 593 GGAATCGGCTTCAAGAGCAAGTGTGCTGTCTGTCTGTGCTCAAGAAACATGACT 652
Db 563 GGAATCGGCTTCAAGAGCAAGTGTGCTGTCTGTGCTGTGCTCAAGAAACATGACT 622
QY 653 TCTTGGTGGAGACCCGAGCTGTGACCAATCTTTGGAGTCCGCGGAGCCATTAATG 712
Db 623 ACTTTGAGAGGAGACCCGAGCTGTGACCAATCTTTGGAGTCCGCGGAGCCATTAATG 682
QY 713 TTTTGAATCTTAACTGTCTCCATGAGCAAGGCTTATTCACAAAGCTCATGAGAGA 772
Db 683 TCTTATCTTCTTATTTATCTCCCTGACCAAGAAATCTCTCATGAGCAATTTCCGAGA 742
QY 773 GTGGGTGGCATATCTCTTAACTGTGAGGCCATGATTTATGAGAAAGTGAAGCTTGC 832
Db 743 GTGGGTGGCTCTCTTTCAGTCTTTCAGAGAAAGAACCAAGTCTTGTGCTAGAAA 802
QY 833 AGGTGTTCACATTTCTGTGTGAACATGAGTGCAGACTGAGAGCCCTGTGAGTGC 892
Db 803 TTGCATGAAAGTGGGTAAACCAACCACTGGCTGTGATGTTCACTGCTGCGCC 862
QY 893 TGAGACAAACCTCCAGAGCTGCTGAC----- 923
Db 863 AGAAGACAGAGAAAGAACTCATGAGGTGACATTTGAATTAATGAGCTGTAGATC 922
QY 924 -----CCTCAGCAGAAACAAAGCTTTCATCAGTGTGATGTGCTTCT 973
Db 923 TAGTTGGCAGACCCAAAGAGAACCGCTTCTGACCACTGTGATGTGAGTGTGCTGC 982
QY 974 TTCTTAAAGAGCTCTGATCTATTTGTCTGAGAAAGATTTAAAGCAATTCCTTCACTA 1033
Db 983 TGCCAAAGCAGCTGACAGATTTCTGACAGAAAGAAATTCACATGCTCCCTACATGG 1042
QY 1034 TCGAGTCAATTAACAGAGATGTGCTTCTGCTGCTTATGAAAGAGCTCTGAGATTC 1093
Db 1043 TGGAAATTAACAGAGATTTGCTGATTTATCCCAATGCAATGTGGCTATTCAC 1102
QY 1094 TCAGTGCTCCCAAGATCCCTT-----GCCCTCACTGATTAACAAACATCC 1141
Db 1103 TCTCTGAAGCAACTGACCAAGACAGCTTACAGAACTTGTGTGAGTCTTACCCCA 1162
QY 1142 TGCATATCCGCTCAAGTATTTGCACTGTGTGCTAATG---AATCTTCATGACAGAC 1198
Db 1163 TTGTCAATGTCTTAAAGAGAGTCACTCAAGTGCACATGAGATTTTAAAGAGGAGCAG 1222
QY 1199 ACTCCCTGAAGTAAATCCAGACAGCTTCTGAGACTTGTGAGATGTGTTTGTG 1258
Db 1223 ATGACCTGTCAAAAGAAAGACTTGTCTGTGACATGTCTGAGATTTGTTATTTGATG 1282
QY 1259 TCCCTGACATGATCAGAGCTGATATCAAGAGATGTGAGCACTGTCTACTTATG 1318
Db 1283 TCCCATCTGTGAATGTGCTGTGACCAAGAGATGTGAGGCCCACTTATATATG 1342
QY 1319 AGTTTGGGACCGGCTCAAGTCTTGAAGACAGAGCCGCTTTTGAAGCCGAGC 1378
Db 1343 AGATGTGATGCCCCAAGCTTCTATCAGACATGAGACCCAAAGCATGATGAGGAGC 1402
QY 1379 ACAGTGAAGTCCGCTTGTGTGTGAGTGTGCTTCTGAAAGGAGGAGCATTTTATGT 1438
Db 1403 ATGAGATGAATCTTCTGTCTTGAAGCCCGCTTTTAAAGAG----- 1449
QY 1439 TCGAAGAGCCACAGAGAGAGAAATTAAGTGAAGTGAATTAATCTGAGCTA 1498
Db 1450 -----GTGCCACAGAAAGAGATCAAACTGAGCAAGTGTGATGAATAATCTGGGCA 1504
QY 1499 CCTTGTCTGAACCGGGAATCTTAATGGAAGAGACTGTCTTGTGCGCAGTTAATC 1558
Db 1505 ACTTTGTGAATGGAATTCATGAGAGGCTTCTCAAGGCGCAGATATGACT 1564
QY 1559 TGAATGAGAGTACTCCAGCTGAGACTTGAACATGAGCTCGGACAGAGACTCAAGAAC 1618
Db 1565 ACAAGGAAGTAACTGTGAGATTTGAGGACCAACCAAGCAGCCAGAACTGAAGACA 1624
QY 1619 CCGGAGTGAATTTTGAAC 1637

Db 1625 AGGAGTGGCTTTCTGAC 1643

RESULT 9

```

US-09-810-861B-5
; Sequence 5, Application US/09810861B
; Patent No. 6770799
; GENERAL INFORMATION:
; APPLICANT: Mori, Tsafir S.
; APPLICANT: Soreq, Hermona
; APPLICANT: Arntzen, Charles J.
; APPLICANT: Maeson, Hugh S.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT HUMAN ACETYLCHOLINESTERASE IN
; TITLE OF INVENTION: TRANSGENIC PLANTS
; FILE REFERENCE: Bt1-45
; CURRENT APPLICATION NUMBER: US/09/810,861B
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/190,440
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 5
; LENGTH: 1725
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: human acetylcholinesterase gene optimized for
; OTHER INFORMATION: expression in plants
; US-09-810-861B-5

```

Query Match 8.0%; Score 139.4; DB 4; Length 1725;

Qy	175	CGTGGAAACGTCCTCTCTGGAGATCCCTCTTGGCTGCTCCCCCGTGGGAATCCCGCATTT	234
Db	175	CTGTCTCTGCTTCTCTGGGACATCCCTCTTGGGAGCACCAATGGGACCCCGTCGCTTT	234
Qy	235	ACGAACCCGACAGCTGCAATCGCCCTGGGATTAACCTTGAGAAAGCCACCTCTCAACCTTAAT	294
Db	235	CTGCCACCCGGAGCCCAAGCAGCTTGTGTACGGGGGTGTAGACGCTACCAACTTCCAGAGT	294
Qy	295	TTGTGCTCTCCAGAACTCAGAGTGGCTGCTCTTAAGATCAACACATGCTCAAGGTGCTTAC	354
Db	295	GTGTGCTACCAATATGTGGACACCCCTATACCAGGTTTGAAGGACACGAGATGTGGAAAC	354
Qy	355	CCGAATTCGAGGTGACAAAGATGCTCTCTACATCAACATCTATGCGCTGCGCCACGCC	414
Db	355	CCCAACCGTGAAGCTAGGAGGACCTGCTGTAACCTCAAGTGTGGACACCTATACCCCGG	414
Qy	415	GATACAGGCTCCAAAGCTCCCGGTCTGGTGGTGGTTCCAGAGAGTACCTTCAAGACTGC	474
Db	415	CCTACATCCCCCA---CCCTGTCTCTGCTGTGATTTATGGGGGTGGCTTTCAAGTGGG	471
Qy	475	TCAGCTC-----CATCTTGTATGGGTCCGCCCTGCTGCTATAGAGACGTGCTGTT	528
Db	472	GCTCTCTCTTGGAGGTGTAGAGATGGCCGCTTCTTGTGTACGGCCGAGAGGACTGTGCTG	531
Qy	529	GTGTCGTCCAGTACCCGGCTAGGAATATTTGGTTCTTCAAC--ACATGGATCAACAT	585
Db	532	GTGTTCATGAACCTACCCGGGTGGAGGCTTTGGCTTCTGAGCCTGTGCCGGGGAGACCCAGAG	591
Qy	586	GCTCCGGGGAACCTGGGCTTCAAGGACCAAGGTGCTGTCTGTCTGGGTCCAGAAAGAAC	645
Db	592	GCCCCGGCAAGTGTAGTCTCTGTGATCAAGGCTGACCTGTGAGTGGGTGACAGAGAAC	651
Qy	646	ATCGAGTTCTTTCGTGGGAGACCCACAGCTCTGTGACATCTTTTGGCAGAGTCCGCGGAGCC	705
Db	652	GTGGACAGCTTTCGGGGGTGACCCCGACATCAATGTAACGCTGTTTGGGGAGAGCGCGGAGCC	711
Qy	706	ATTAAGTGTTCATAGTCTTAATCTGTCTCCCATGCGCAAGAGGCTTATTTCCAGAAAGCATC	765

Db 712 GCCTCGGTGGGCATGCACCTGCTGTCCCCCGCCAGCCGGGGCTGTTCACAGGGCCGTG 771

QY 766 ATGAGAGTGGCGTGCCCA 784
|||
Db 772 CTGCAGAGCGGTGCCCA 790

RESULT 10
US-07-732

US-07-732-962A-1
Sequence 1, Application US/07732962A
Patent No. 5248604
GENERAL INFORMATION:
APPLICANT: Fischer, Mel
TITLE OF INVENTION: EXPRESSION OF ENZYMATICALLY ACTIVE
TITLE OF INVENTION: RECOMBINANT HUMAN ACETYLCHOLINESTERASE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White, Esq.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/732,962A
FILING DATE: 19910722
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39304/JPW/LSW
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UT
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
US-07-732-962A-1

Query Match 8.0%; Score 139.4; DB 1; Length 1845;

Qy	175	CCGTGGAAGTGTCTCTGGAAATCCCTTTGTGTCGCCCGCTGGAAATCCCGGAAATT	234
Db	175	CCGTCTCTGCTTCTCTGGCAATCCCTTTGGGAGGCCACCAATGGGACCCCGTGCTTT	234
Qy	235	ACGAACCCGACAGCTTCATCGCCCTGGGATATCTTGACGAAGCACCTCTCAACCTTAAT	294
Db	235	CTGCCACCGGAGGCCAAGCAGCTTTGGTCAAGGGGTGTGAGCGCTACAACTTCAAGT	294
Qy	295	TTGTGCTTCAAGAACTCAGAGTGCTGCTCTTAATCAACATGCTCAAGGTSCATTAC	354
Db	295	GTCTGCTAACCAATATGTGGACACCTTAACCAAGGTTTGGGGCAACCGAGATGTGAAAC	354
Qy	355	CCGAATTTGGAGTGTCAAGAAACATGCTCTTAACCTGAACATCTATGGCGCTGGCCAGGC	414
Db	355	CCCAACCGGTAGCTGAGCGAGGATCGCTGTAACTCAACGTTGTGACACCAATACCCCGG	414

415 GATACAGGCTCCAGAGCTCCCGCTTGTGTGCTCCAGAGAGTCCCTTCAAGACTGGC 474
415 CCTACATCCCA---CCCTGTCTCTGTGATCTAAGGGGTGCTTCAAGTGG 471
415 TGAAGCTC-----CATCTTGTATGGGTCCGCTGTGCTTATAGAGAGTGTGTT 528
412 GCTCTCTCTTGGAGCTGTAGCATGGCCCTTCTTGTACAGGCCAGAGAGACTGTGCTG 531
529 GTGTGTCTCAGATACCGGCTAGGAATTTGGTTCTTACAC---ACATGGATCAGCAT 585
532 GTGTCTCATATCACTACCGGCTGGAGCTTGTGCTTCTGCTGTGCGGGAGCCGAGAG 591
586 GCTCCGGGAGACTGGGCTTCAAGGACGAGTGTGTCTGTCTGTGCTGTGCTTCAAGAGAAC 645
592 GCGCCGGGAGATGTGGGTCTCTGGATCAGAGGCTGGCCCTGCAATGGGTGAGAGAAC 651
646 ATCGAGTTCTTGTGTGGGAGACCCAGCTGTGTGACATTTTGGGAGTCCGCGGAGCC 705
652 GTGGCAGCTTGTGGGGGTGACCCGACATCATGTGACGCTTTTGGGGAGAGCGCGGAGCC 711
706 ATAACTTTTCTAATCTTACTGTCTCCATGGCCAAAGCTTATCCAAAGCCATC 765
712 GCTCTGTGTGGGATGACTGTCTGTCTCCGCGGAGCGGGGCTGTCCACAGGACCGTG 771
766 ATGAGAGTGGGGTGGCCA 784
772 CTGACAGCGGTGGCCCA 790

RESULT 11

PCT-US92-06106-1
Sequence 1, Application PC/TUS9206106
GENERAL INFORMATION:
APPLICANT: Flascher, Meli
TITLE OF INVENTION: ENZYMATICALLY ACTIVE RECOMBINANT HUMAN
NUMBER OF SEQUENCES: 2
TITLE OF INVENTION: ACETYLCHOLINESTERASE AND USES THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White, Esq.
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06106
FILING DATE: 19920722
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39304-B-PCT/JPW/EAB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1845 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1842
PCT-US92-06106-1

Query Match 8.0%; Score 139.4; DB 5; Length 1845;
Best Local Similarity 55.1%; Pred. No. Se-33;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

175 CTTGTGAACGTGTCTCTCGAGATCCCTTTGTGCTCCCGCTGGAGATCCCTGCAATT 234
175 CTTGTCTGTGCTTCTCTGGGATCCCTTTGGAGAGCCACCAATGGAGACCCCTGTGCTTT 234
225 ACGAAGCCGACAGCTGCAATCCGCTTGGATTAATTGGAGAGACCACTCTTCACTTAAT 294
225 CTGCAACCGGAGCCCAAGACAGCTTGTGTCAAGGGGTGTAGACGCTTCAACCTTCCAGAGT 294
225 TTGTGCTTCCAGAACTCAGAGTGGCTGCTTATAGATCAACATGCTCAAGGTGATTAAC 354
225 GTCTGTCAATATATGTGAGAACCTTATACCAAGTTTGTAGGGGACCTGAGATGTGAAAC 354
355 CGGAATTCGAGATGTCAAGAACTGCTCTTACCTGAAATCTATATGCGCTCCAGCCAGCC 414
355 CCAACCGTGAAGTGAAGGAGAGATGCTGTATACCTCAACGTGTGAGACATACCCCGG 414
415 GATACAGGCTCCAGACTCCCGCTTGTGTGTGTGTCCAGAGAGTCCCTTCAAGACTGGC 474
415 CCTACATCCCA---CCCTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 471
475 TGAAGCTC-----CATCTTGTATGGGTCCGCTGTGCTGTGCTGTATGAGAGAGTGTG 528
472 GCTCTCTCTTGTGAGAGTGTATGATGCGGCTTCTTGTGTATGAGGCGGAGAGACTGTG 531
529 GTGTGTCTCAGTACCGGCTAGGAATTTGGTTCTTACAC---ACATGGATCAGCAT 585
532 GTGTCTCATGAACTACCGGCTGGAGACCTTGTGGCTTCTTGGCCCTGTGCGGGGAGCGAGAG 591
586 GCTCCGGGAGACTGGGCTTCAAGAGACAGTGTGTCTGTGTCTGTGTGTGTGTGTGTGTGT 645
592 GCGCCGGGCAATGTGGGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 651
646 ATCGAGTTCTTGTGTGGGAGACCCAGCTGTGTGACATCTTGTGCGAGTCCGCGGAGCC 705
652 GTGGCAGCTTGTGGGGGTGAGCCGACATCATGTGACGCTTTTGGGGAGAGCGCGGAGCC 711
706 ATAACTTTTCTAATCTTACTGTCTCCATGGCCAAAGCTTATCCAAAGCCATC 765
712 GCTCTGTGTGGGATGACTGTCTGTCTCCGCGGAGCGGGGCTGTCCACAGGACCGTG 771
766 ATGAGAGTGGGGTGGCCA 784
772 CTGACAGCGGTGGCCCA 790

RESULT 12

US-09-949-016-1192
Sequence 1192, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C1001307
CURRENT APPLICATION NUMBER: US/09/949,016
FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1192
LENGTH: 2158
TYPE: DNA
ORGANISM: Human
US-09-949-016-1192

Query Match 8.0%; Score 139.4; DB 4; Length 2158;
Best Local Similarity 55.1%; Pred. No. 5.6e-33;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

```
175 CCTGTGAACGTTGCTCTGAGATCCCTTTGCTGCTCCCGCTGGAGATCCCTGGCATTT 234
269 CCGTCTCTGCTTCTCTGGGATCCCTTTGCGAGCCACCATGGAGACCCCGTGGCTTT 328
225 ACGAACCCGACGCTGATCGCCCTGGGATTAATCTTGGAGAACCACTCTCACTTAAT 294
329 CTGCCACCGAGCCCAAGACGCTTGTCTAGAGGGGTGATGACGTACCAACTTTCAGAT 388
225 TTGTCCTCCAGAACTCAGAGTGGCTGCTTAGATCAACATGCTCAAGGTGATTAAC 354
389 GTCTCTACCAATATGTGACACCTTATACCAAGTTTGAAGGACACGAGATGTGAAAC 448
355 CCGAATTCGAGTGTCAAGAGACTGCTCTAATCACTGAACATCTATGCGCTGCCAGCC 414
449 CCGAACGCTGAGCTAGAGGAGACTGCTGTACTCAAGTGTGAGACACATACCCCGG 508
415 GATACAGCTCCAGCTCCCGCTTGTGTGTGCTTCCAGAGTGGCTTCAAGACTGAC 474
509 CCTACATCCCCCA---CCCTGTCTCTGTGATCTATGAGGGGTGGCTTCTACAGTGG 565
475 TCAGCTC-----CATCTTGATGGGTCCGCTGCTGCTGCTATGAGGACGTGCTGTT 528
566 GCTCTCTCTTGAAGCTGATGACATGACGCTTCTTGTGATCAGGCGGAGAGACTGTG 625
529 GTGTGCTGTCAAGTACCGGCTAGAGAAATTTGGTTTCTTCAAC---ACATGGATCAGCAT 585
626 GTGTCTCAATGAACTACCGGGTGGAGCTTTGGCTTCTGAGCTTCCCGGGAGCCGAGAG 685
586 GCTCCGGGAACTGGGCTTCAAGGACAGGTGCTCTGTCTCTGGTCTCAAGAAAC 645
686 GCCCGGGCAATGTGGGTCTCTGATCAGAGGCTGCGCTCAGTGGGTGACAGGAAAC 745
646 ATCGAGTCTTCTGGTGGGAGCCCGAGCTCTGACATCTTTGGGAGTCCCGGGAGCC 705
746 GTGGAGCTTCTGGGGGTGACCCGACATGATGAGCTGTTTGGGAGAGCCGGGAGCC 805
706 ATAAGTGTCTTCTGATCTTATCTGCTCCCATGGCCAAAGCTTATCCAAAGCATC 765
806 GCCTCGGTGGGATCAGCTGCTGTCTCCCGCCAGCCGGGGCTGTTCACAGGGCCGTG 865
766 ATGAGAGTGGGCTGAGCA 784
866 CTGACAGCGGTGCCCCCA 884
```

RESULT 13
US-09-949-016-1193
; Sequence 1193, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1193
; LENGTH: 2158
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-1193

Query Match 8.0%; Score 139.4; DB 4; Length 2158;
Best Local Similarity 55.1%; Pred. No. 5.6e-33;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

```
175 CCTGTGAACGTTGCTCTGAGATCCCTTTGCTGCTCCCGCTGGAGATCCCTGGCATTT 234
269 CCGTCTCTGCTTCTCTGGGATCCCTTTGCGAGCCACCATGGAGACCCCGTGGCTTT 328
225 ACGAACCCGACGCTGATCGCCCTGGGATTAATCTTGGAGAACCACTCTCACTTAAT 294
329 CTGCCACCGAGCCCAAGACGCTTGTCTAGAGGGGTGATGACGTACCAACTTTCAGAT 388
225 TTGTCCTCCAGAACTCAGAGTGGCTGCTTAGATCAACATGCTCAAGGTGATTAAC 354
389 GTCTCTACCAATATGTGACACCTTATACCAAGTTTGAAGGACACGAGATGTGAAAC 448
355 CCGAATTCGAGTGTCAAGAGACTGCTCTAATCACTGAACATCTATGCGCTGCCAGCC 414
449 CCGAACGCTGAGCTAGAGGAGACTGCTGTACTCAAGTGTGAGACACATACCCCGG 508
415 GATACAGCTCCAGCTCCCGCTTGTGTGTGCTTCCAGAGTGGCTTCAAGACTGAC 474
509 CCTACATCCCCCA---CCCTGTCTCTGTGATCTATGAGGGGTGGCTTCTACAGTGG 565
475 TCAGCTC-----CATCTTGATGGGTCCGCTGCTGCTGCTATGAGGACGTGCTGTT 528
566 GCTCTCTCTTGAAGCTGATGACATGACGCTTCTTGTGATCAGGCGGAGAGACTGTG 625
529 GTGTGCTGTCAAGTACCGGCTAGAGAAATTTGGTTTCTTCAAC---ACATGGATCAGCAT 585
626 GTGTCTCAATGAACTACCGGGTGGAGCTTTGGCTTCTGAGCTTCCCGGGAGCCGAGAG 685
586 GCTCCGGGAACTGGGCTTCAAGGACAGGTGCTCTGTCTCTGGTCTCAAGAAAC 645
686 GCCCGGGCAATGTGGGTCTCTGATCAGAGGCTGCGCTCAGTGGGTGACAGGAAAC 745
646 ATCGAGTCTTCTGGTGGGAGCCCGAGCTCTGACATCTTTGGGAGTCCCGGGAGCC 705
746 GTGGAGCTTCTGGGGGTGACCCGACATGATGAGCTGTTTGGGAGAGCCGGGAGCC 805
706 ATAAGTGTCTTCTGATCTTATCTGCTCCCATGGCCAAAGCTTATCCAAAGCATC 765
806 GCCTCGGTGGGATCAGCTGCTGTCTCCCGCCAGCCGGGGCTGTTCACAGGGCCGTG 865
766 ATGAGAGTGGGCTGAGCA 784
866 CTGACAGCGGTGCCCCCA 884
```

RESULT 14
US-08-318-826A-5
; Sequence 5, Application US/08318826A
; Patent No. 5891725
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Eckslein, Fritz
; TITLE OF INVENTION: Synthetic Antisense
; TITLE OF INVENTION: Oligodeoxynucleotides and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Kohn & Associates
; STREET: 30500 No. 5891725thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,826A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 2391.00001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (248) 539-5050
TELEFAX: (248) 539-5050
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2256 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: /note="Splice variant: Exons 1, 2,
OTHER INFORMATION: 3, 4 and 6"
US-08-318-826A-5

Query Match 8.0%; Score 139.4; DB 2; Length 2256;
Best Local Similarity 55.1%; Pred. No. 5.8e-33;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;
175 CCTGGAAGTGTCTGCGAGTCCCTTGTGCTCCCGCGTGGATCCCTGGATT 234
334 CTTGCTCTGCTTCTTGGCATCCCTTGGAGCCACCATGGACCCGCTGCTT 393
235 ACGAACCCGACGCTGATCGCTGGATTAATTGGAGAACCATCTCTACCTAAT 294
394 CTGCCACCGAGCCCAAGAGCTTGGTCAAGGGGTGATGACGCTACAACTTCAAGT 453
295 TTGTGCTCCAGAACTCAGAGTGTCTCTTAATCAACACATCTCAAGTGATTAC 354
454 GTGTGCTACCAATATGATGACACCTATACCAAGTTTGGAGGACCGAGATGTGAAC 513
355 CCGAATTCGAGAGTCAAGAACTGCTCTACCTGAACATCTATAGCGCTGCCACGCC 414
514 CCGAACCTGAGCTGAGGAGGAGTGTCTGTAACCTCAACGTGTGACACCATACCCCG 573
415 GATACAGGCTCAAGCTCCCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 474
574 CTTACATCCCCCA---CCCTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 630
475 TCAGCTC-----CATCTTGAATGGGTCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 528
631 GCTTCTCTCTTGAAGGTATGATGAGTCCGCTTCTTGTGAACGCGCGAAGACGTGTGT 690
529 GTGTGTCTCAGTACCGGCTAGAAATATTTGTTCTTCAAC---ACATGGATCAGCAT 585
691 GTGTGCTATGATCTACCGGCTGAGGACCTTGTGCTTCTGTGTGTGTGTGTGTGTGT 750
586 GCTTCCGAGAACTGAGGCTTCAAGAACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 645
751 GCCCGGCGCAATGTGTGTCTCTGTGATCAGAGCTGTGTGTGTGTGTGTGTGTGTGT 810
646 ATGAGTTCTTGGTGGGAGCCCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 705
811 GTGCAAGCTTCTGGGGGTGACCCCAATCATGTAGCGCTTGTGGAGAGGCGGAGGCC 870
706 ATAGTGTCTTACTGTATATCTGTCTTCCATGGCCAAAGCTTATTTCAAAAGCCATC 765
871 GCCTGGGTGGGCAATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 930

QY 766 ATGAGAGTGGGTGGCCCA 784
DB 931 CTGCAAGCGGTGCCCCCA 949

RESULT 15
US-08-370-156-1
Sequence 1, Application US/08370156
Patent No. 5932780
GENERAL INFORMATION:
APPLICANT: Soreq, Hermona
APPLICANT: Zakut, Haim
APPLICANT: Shani, Moshe
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Reising, Ethelington, Barnard & Perry
STREET: P. O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
ZIP: 48099
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,156
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-307 (Mullford)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2256 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-370-156-1

Query Match 8.0%; Score 139.4; DB 2; Length 2256;
Best Local Similarity 55.1%; Pred. No. 5.8e-33;
Matches 341; Conservative 0; Mismatches 266; Indels 12; Gaps 3;

QY 175 CCTGGAAGTGTCTGCGAGTCCCTTGTGCTCCCGCGTGGATCCCTGGATT 234
DB 334 CTTGCTCTGCTTCTTGGCATCCCTTGGAGCCACCATGGACCCGCTGCTT 393
QY 235 ACGAACCCGACGCTGATCGCTGGATTAATTGGAGAACCATCTCTACCTAAT 294
DB 394 CTGCCACCGAGCCCAAGAGCTTGGTCAAGGGGTGATGACGCTACAACTTCAAGT 453
QY 295 TTGTGCTCCAGAACTCAGAGTGTCTCTTAATCAACACATCTCAAGTGATTAC 354
DB 454 GTGTGCTACCAATATGATGACACCTATACCAAGTTTGGAGGACCGAGATGTGAAC 513
QY 355 CCGAATTCGAGAGTCAAGAACTGCTCTACCTGAACATCTATAGCGCTGCCACGCC 414
DB 514 CCGAACCTGAGCTGAGGAGGAGTGTCTGTAACCTCAACGTGTGACACCATACCCCG 573
QY 415 GATACAGGCTCAAGCTCCCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 474
DB 574 CTTACATCCCCCA---CCCTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 630
QY 475 TCAGCTC-----CATCTTGAATGGGTCCGCTGTGTGTGTGTGTGTGTGTGTGTGT 528

```
Db 631 GCCTCTCTCTTGACGTGACGATGCGCCTTCTTGATACAGCCGAGAGACTGTGCTG 690
Oy 529 GTGTCTCTCACTACCGGCTAGGAATTTGGTTCTTCACC--ACATGGATCAGCAT 585
Db 691 GTGTCAATGAACTACCGGGGTGAGCCTTGGCTTCTGCCCCGCGGGAGCCGAGAG 750
Oy 586 GCTCCGGGGAACCTGGGCTTCAAGGACGAGTGGCTGTCTGCTGGGTCCAGAGAAG 645
Db 751 GCCCGGGCAATGTGGGTCTCTGATCAGAGGCTGGCCCTGCAGTGGGTGCAGAGAAC 810
Oy 646 ATCGAGTTCTTGGGTGGGACCCCACTCTGTGACCAATCTTGGCGAGTCCGCGGAGCC 705
Db 811 GTGCGAGCCTTCCGGGGGTGACCCGACATCATGACGCTGTTGGGGAGAGCGGGAGGCC 870
Oy 706 ATAGTGTCTTCTAATCTGCTGCTCCCATGSCCAAGGCTTATCCAGAAAGCCATC 765
Db 871 GCCTGGGTGGGCAATGCACTGTCTGTCGCCGCCAGCCGGGGGCTGTTCACAGGGCCGTTG 930
Oy 766 ATGAGAGTGGGGTGGCCA 784
Db 931 CTGCAAGCGGTGCCCCCA 949
```

Search completed: June 14, 2005, 23:03:36
Job time : 296.702 secs

This Page Blank (uspto)

PT surrogate markers, in tissue typing and chromosome mapping.
XX
PS Claim 1, Page 109-110, 122pp; English.
XX

CC The present nucleotide sequence, the coding region of which is also
CC claimed, is that of cDNA encoding human 53010, a novel member of the
CC carboxylesterase family. The invention provides 53010 nucleic acids,
CC antisense molecules, expression vectors, host cells, transgenic animals,
CC 53010 proteins, fusion proteins, antigenic peptides, anti-53010
CC antibodies and methods for detecting the presence of 53010 polypeptides
CC or nucleic acids, of identifying a compound that binds to the 53010
CC polypeptide, and of modulating the activity of the polypeptide. The 53010
CC nucleic acids and polypeptides can act as novel diagnostic and
CC therapeutic agents for controlling disorders involving aberrant or
CC deficient hydrolysis of carboxylic esters. As 53010 mRNA is highly
CC expressed in the central and peripheral nervous system, and its
CC expression is regulated in some rodent pain models, 53010 molecules can
CC also act as novel diagnostic targets and therapeutic agents for
CC controlling neurological disorders, such as pain-related disorders. A
CC claimed method of treating or preventing a disorder (especially a pain-
CC related disorder) characterised by aberrant activity of a 53010-
CC expressing cell involves administering a compound that modulates 53010
CC activity or expression. 53010 nucleic acids are also useful in chromosome
CC mapping, tissue typing, in forensic biology, prognostic assays, in
CC arrays, for detection of variations or mutations, as surrogate markers
CC and in pharmacogenomics
XX

Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;

Query Match 100.0%; Score 1746; DB 6; Length 2158;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCACAGGAGCTTACTCATCTGCTTCAATGAGTGTCTTTCTTCAATTCCTCAAGCC 60
DB 96 ATGCCACAGGAGCTTACTCATCTGCTTCAATGAGTGTCTTTCTTCAATTCCTCAAGCC 155
QY 61 CTGTTGGGACACAGACAGTGGGAAAAAATGGGCCCTTCTGTGAAGGGCCACAGAGAAC 120
DB 156 CTGTTGGGACACAGACAGTGGGAAAAAATGGGCCCTTCTGTGAAGGGCCACAGAGAAC 215
QY 121 ACCAGAGCTGGATGATTCAGAGGCAAGCATCTGTCTGGGAAAGCCCTGTGCTGTG 180
DB 216 ACCAGAGCTGGATGATTCAGAGGCAAGCATCTGTCTGGGAAAGCCCTGTGCTGTG 275
QY 181 AACGTGTTCTCGAGTCCCTTTGTCTGCTCCCGCTGGAGTCCCTGCGATTATTCGAA 240
DB 276 AACGTGTTCTCGAGTCCCTTTGTCTGCTCCCGCTGGAGTCCCTGCGATTATTCGAA 335
QY 241 CCGCAGCTGCATCGCCCTGGGATTAATTGGAGAAAGCACCTCTCAATTATTTGTGC 300
DB 336 CCGCAGCTGCATCGCCCTGGGATTAATTGGAGAAAGCACCTCTCAATTATTTGTGC 395
QY 301 CTCCGAACTCAGAGTGGCTGCTTAGTCAACAATGCTCAAGAGTGCATTAATCCGAAA 360
DB 396 CTCCGAACTCAGAGTGGCTGCTTAGTCAACAATGCTCAAGAGTGCATTAATCCGAAA 455
QY 361 TTTCGAGTGTCAAGAACTGCTTACTGAACATCTATGCGCTGTGCCACGCGGATACA 420
DB 456 TTTCGAGTGTCAAGAACTGCTTACTGAACATCTATGCGCTGTGCCACGCGGATACA 515
QY 421 GGCTCCAAAGCTCCCGCTTTGGTGTGTCTCCAGAGAGGTGCTTCAAGACTGCTCAGCC 480
DB 516 GGCTCCAAAGCTCCCGCTTTGGTGTGTCTCCAGAGAGGTGCTTCAAGACTGCTCAGCC 575
QY 481 TCCATCTTTGATGGGTCCGCTGCTGCTATGAGGAGCGTGTGTGTGTGTGTGTGTGTGT 540
DB 576 TCCATCTTTGATGGGTCCGCTGCTGCTATGAGGAGCGTGTGTGTGTGTGTGTGTGTGT 635
QY 541 TACCGGCTAGGAATATTTGTTCTTCAACCATGGAATCAGCATGCTCCGGGAACTGG 600
DB 636 TACCGGCTAGGAATATTTGTTCTTCAACCATGGAATCAGCATGCTCCGGGAACTGG 695

QY 601 GCTTCAAGACCAAGAGTGGCTGCTGTCTGAGGTCAGAAAGATCGAGTTCTTCGCT 660
DB 696 GCTTCAAGACCAAGAGTGGCTGCTGTCTGAGGTCAGAAAGATCGAGTTCTTCGCT 755
QY 661 GGGAGCCCAAGCTCTGTGACCAATCTTTGGCCAGATCCCGGGAGCCATAGTGTCTTGT 720
DB 756 GGGAGCCCAAGCTCTGTGACCAATCTTTGGCCAGATCCCGGGAGCCATAGTGTCTTGT 815
QY 721 CTATATCTGTCTCCATGAGCCCAAGAGCTTATTCACAAAGCATGAGAGTGGAGTG 780
DB 816 CTATATCTGTCTCCATGAGCCCAAGAGCTTATTCACAAAGCATGAGAGTGGAGTG 875
QY 781 GGCATCATCCCTTACCTGAGAGCCCATATTTATGAGAAAGTGAAGACCTTCAGTGGTT 840
DB 876 GGCATCATCCCTTACCTGAGAGCCCATATTTATGAGAAAGTGAAGACCTTCAGTGGTT 935
QY 841 GCACATTTCTGTGTGTAATATGCGTCAGACTGTGAGGCCCTGAGAGTGTCTGAGACA 900
DB 936 GCACATTTCTGTGTGTAATATGCGTCAGACTGTGAGGCCCTGAGAGTGTCTGAGACA 995
QY 901 AAACCTTCAAGAGCTGTGACCTTCAGCCAGCAAAACAAAGTCTTTCATCTGAGTGGTT 960
DB 996 AAACCTTCAAGAGCTGTGACCTTCAGCCAGCAAAACAAAGTCTTTCATCTGAGTGGTT 1055
QY 961 GATGGTCTTCTTCTTCAATGAGCCCTTATGATCTATTTGTCTGAGAAAGCTTTAAAGCA 1020
DB 1056 GATGGTCTTCTTCTTCAATGAGCCCTTATGATCTATTTGTCTGAGAAAGCTTTAAAGCA 1115
QY 1021 ATTCTTCATCATCTGAGAGTCAATTAACACAGAGTGTGGCTTCTGCTGCTATGAGAG 1080
DB 1116 ATTCTTCATCATCTGAGAGTCAATTAACACAGAGTGTGGCTTCTGCTGCTATGAGAG 1175
QY 1081 GCTCTGAGATCTCTCATGAGTCTCAACAAGTCCCTTCCATCTGATACAAACATC 1140
DB 1176 GCTCTGAGATCTCTCATGAGTCTCAACAAGTCCCTTCCATCTGATACAAACATC 1235
QY 1141 CTGCACATCCCGCTCAGTATTTGGACCTTGTGGCTTAATATCTTCATGCAAGAC 1200
DB 1236 CTGCACATCCCGCTCAGTATTTGGACCTTGTGGCTTAATATCTTCATGCAAGAC 1295
QY 1201 TCCCTGACTGAATCCGAGACAGTCTTCTGAGCTTGTGGAGATGTCTTCTTGTGCT 1260
DB 1296 TCCCTGACTGAATCCGAGACAGTCTTCTGAGCTTGTGGAGATGTCTTCTTGTGCT 1355
QY 1261 CTTGCACTGATCAAGTCTGATATCAACAGAGTGTGTGACACTGTCTATCTTATGAG 1320
DB 1356 CTTGCACTGATCAAGTCTGATATCAACAGAGTGTGTGACACTGTCTATCTTATGAG 1415
QY 1321 TTTCGGACCGGCTCAGTCTTGTGAAGACAGAAAGCCGCTTTGTCAAAAGCCGACAC 1380
DB 1416 TTTCGGACCGGCTCAGTCTTGTGAAGACAGAAAGCCGCTTTGTCAAAAGCCGACAC 1475
QY 1381 GCTGATGAATCCGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
DB 1476 GCTGATGAATCCGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1535
QY 1441 GAAGAGCCCAAGAGAGGAGAAATTACTGAGCCGGAAGATGATGAAATATCGGCTTAC 1500
DB 1536 GAAGAGCCCAAGAGAGGAGAAATTACTGAGCCGGAAGATGATGAAATATCGGCTTAC 1595
QY 1501 TTTCGCTGAACCGGAAATCTTAATGGAAAGCACTGTCTGTGTGCACTTAATATCTG 1560
DB 1596 TTTCGCTGAACCGGAAATCTTAATGGAAAGCACTGTCTGTGTGCACTTAATATCTG 1655
QY 1561 ACTGAGCAATTAATCTCAAGCTGAGCTTGAACATGAGGCTTGGACAGAGACTCAAGAAC 1620
DB 1656 ACTGAGCAATTAATCTCAAGCTGAGCTTGAACATGAGGCTTGGACAGAGACTCAAGAAC 1715
QY 1621 CCGGTGAGATTTTGTGACAGCAACATCCCGTGAATCTGTCTGCGCTCCGACATGCTCAC 1680
DB 1716 CCGGTGAGATTTTGTGACAGCAACATCCCGTGAATCTGTCTGCGCTCCGACATGCTCAC 1775
QY 1681 AGTCTCTTTCTTCTTAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740

DB 1776 ACTCCCTTCTTCTTACTTCTCTCTCTCCACGCTTCTTTCTTTGTGCT 1835
 QY 1741 CCTGA 1746
 DB 1836 CCTGA 1841

RESULT 2
 ADQ89093
 ID ADQ89093 standard; cDNA; 2158 BP.

AC ADQ89093;
 DT 21-OCT-2004 (first entry)

DE Human urological disorder related protein 53010 encoding cDNA SEQ:45.

XX urological disorder; uropathic; cytostatic; urinary incontinence;
 KM benign prostatic hyperplasia; human; gene; 88.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH 96.1841
 FT CDS /*tag= a
 PT /product= "urological disorder related protein 53010"

XX WO2004065576-A2.

XX 05-AUG-2004.

XX 14-JAN-2004; 2004MO-US000750.

XX 15-JAN-2003; 2003US-0440318P.

XX 04-FEB-2003; 2003US-0444783P.

XX 27-MAR-2003; 2003US-0457901P.

XX 08-MAY-2003; 2003US-0468775P.

XX 19-MAY-2003; 2003US-0471614P.

XX 16-JUN-2003; 2003US-0478742P.

XX 18-JUL-2003; 2003US-0488529P.

XX 30-JUL-2003; 2003US-0491156P.

XX 02-SEP-2003; 2003US-0495949P.

XX 26-SEP-2003; 2003US-0506332P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Karicheti V, Siles-Santiago I, Bilsaof SD;

XX MPI: 2004-562167/54.

XX P-PSDB; ADQ89094.

XX Use of polypeptides related to urological disorders, e.g. 44390, 54181,

XX disorder or identifying a compound capable of treating a urological

XX disorder.

XX Claim 1; SEQ ID NO 45; 542bp; English.

CC The present invention describes the use of polypeptides related to
 CC urological disorders for identifying a compound capable of treating a
 CC urological disorder. Identifying a subject having a urological disorder,
 CC or treating a subject having a urological disorder. Also described: (1) a
 CC method for identifying a compound capable of treating a urological
 CC disorder; (2) a method for identifying a subject having a urological
 CC disorder; and (3) a method for treating a subject having a urological
 CC disorder. The compound has uropathic and cytostatic activities. The
 CC polypeptides related to urological disorders are useful for identifying a
 CC compound capable of treating a urological disorder. Identifying a subject
 CC having a urological disorder, or treating a subject having a urological
 CC disorder. Disorders include urinary incontinence and benign prostatic
 CC hyperplasia. The present sequence encodes a human urological disorder
 CC related protein, which is used in the exemplification of the present

CC invention.
 XX Sequence 2158 BP; 515 A; 559 C; 509 G; 575 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 1746; DB 13; Length 2158;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCAGGAGGACTTACTTCACTGCTTCAACAATGATGCTTTTCTGATTCAGGCC 60
 DB ATGCCAGGAGGACTTACTTCACTGCTTCAACAATGATGCTTTTCTGATTCAGGCC 155
 QY 61 CTGTTGGGACACAGACAGTGGGAAAACCTGGCCCTTCTGCTGAAGGACACAGAGAAC 120
 DB CTGTTGGGACACAGACAGTGGGAAAACCTGGCCCTTCTGCTGAAGGACACAGAGAAC 215
 QY 121 ACCAGGCTGGGATGATTCAGGGCAAGCAATGCTGCTGCTGGAAAGCCCTGCTGTG 180
 DB ACCAGGCTGGGATGATTCAGGGCAAGCAATGCTGCTGCTGGAAAGCCCTGCTGTG 275
 QY 216 ACCAGGCTGGGATGATTCAGGGCAAGCAATGCTGCTGCTGGAAAGCCCTGCTGTG 275
 DB ACCAGGCTGGGATGATTCAGGGCAAGCAATGCTGCTGCTGGAAAGCCCTGCTGTG 275
 QY 181 AACGTTCTCTGGAGTCCCTTGTGCTGCTCCCGGCGGATCCCTGATTTACGAAAC 240
 DB AACGTTCTCTGGAGTCCCTTGTGCTGCTCCCGGCGGATCCCTGATTTACGAAAC 335
 QY 276 AACGTTCTCTGGAGTCCCTTGTGCTGCTCCCGGCGGATCCCTGATTTACGAAAC 335
 DB AACGTTCTCTGGAGTCCCTTGTGCTGCTCCCGGCGGATCCCTGATTTACGAAAC 335
 QY 241 CCGCAGCCTGATCCGCTGGGATTAATTCGAGAGACCACTCTTAATTTGTGC 300
 DB CCGCAGCCTGATCCGCTGGGATTAATTCGAGAGACCACTCTTAATTTGTGC 395
 QY 336 CCGCAGCCTGATCCGCTGGGATTAATTCGAGAGACCACTCTTAATTTGTGC 395
 DB CCGCAGCCTGATCCGCTGGGATTAATTCGAGAGACCACTCTTAATTTGTGC 395
 QY 301 CTCAGAACTCAGAGTGGCTGCTTTAGATCAACATGCTCAAGTGATTAACCGAAA 360
 DB CTCAGAACTCAGAGTGGCTGCTTTAGATCAACATGCTCAAGTGATTAACCGAAA 455
 QY 396 CTCAGAACTCAGAGTGGCTGCTTTAGATCAACATGCTCAAGTGATTAACCGAAA 455
 DB CTCAGAACTCAGAGTGGCTGCTTTAGATCAACATGCTCAAGTGATTAACCGAAA 455
 QY 361 TTGGAGTGTCAAGAGTCCCTTAACCTGAACATTAATGAGGCTCCGACGCGATACA 420
 DB TTGGAGTGTCAAGAGTCCCTTAACCTGAACATTAATGAGGCTCCGACGCGATACA 515
 QY 456 TTGGAGTGTCAAGAGTCCCTTAACCTGAACATTAATGAGGCTCCGACGCGATACA 515
 DB TTGGAGTGTCAAGAGTCCCTTAACCTGAACATTAATGAGGCTCCGACGCGATACA 515
 QY 421 GGCTCCAGGCTCCCGCTGCTGCTGCTGCTTCCAGAGAGTCCCTTAAGTGGCTCAGCC 480
 DB GGCTCCAGGCTCCCGCTGCTGCTGCTGCTTCCAGAGAGTCCCTTAAGTGGCTCAGCC 575
 QY 516 GGCTCCAGGCTCCCGCTGCTGCTGCTGCTTCCAGAGAGTCCCTTAAGTGGCTCAGCC 575
 DB GGCTCCAGGCTCCCGCTGCTGCTGCTGCTTCCAGAGAGTCCCTTAAGTGGCTCAGCC 575
 QY 481 TCCATCTTTGATGGTCCGCTGCTGCTGCTTATGAGAGAGTCTGCTGCTCAG 540
 DB TCCATCTTTGATGGTCCGCTGCTGCTGCTTATGAGAGAGTCTGCTGCTCAG 635
 QY 576 TCCATCTTTGATGGTCCGCTGCTGCTGCTTATGAGAGAGTCTGCTGCTCAG 635
 DB TCCATCTTTGATGGTCCGCTGCTGCTGCTTATGAGAGAGTCTGCTGCTCAG 635
 QY 541 TACCGGCTAGAAATATTTGCTTTCTTCCACATAGGATCAAGATGCTCCGGGAACTGG 600
 DB TACCGGCTAGAAATATTTGCTTTCTTCCACATAGGATCAAGATGCTCCGGGAACTGG 695
 QY 636 TACCGGCTAGAAATATTTGCTTTCTTCCACATAGGATCAAGATGCTCCGGGAACTGG 695
 DB TACCGGCTAGAAATATTTGCTTTCTTCCACATAGGATCAAGATGCTCCGGGAACTGG 695
 QY 601 GCCTTCAAGAGACAGAGTGGCTGCTGCTGCTGCTTCCAGAGAGTCCGATGCTTCCGCT 660
 DB GCCTTCAAGAGACAGAGTGGCTGCTGCTGCTGCTTCCAGAGAGTCCGATGCTTCCGCT 755
 QY 696 GCCTTCAAGAGACAGAGTGGCTGCTGCTGCTGCTTCCAGAGAGTCCGATGCTTCCGCT 755
 DB GCCTTCAAGAGACAGAGTGGCTGCTGCTGCTGCTTCCAGAGAGTCCGATGCTTCCGCT 755
 QY 661 GGGGACCCCAAGCTCTGTGACATCTTTGGAGTCCGGGAGCCATTAAGTCTTCTAGT 720
 DB GGGGACCCCAAGCTCTGTGACATCTTTGGAGTCCGGGAGCCATTAAGTCTTCTAGT 815
 QY 756 GGGGACCCCAAGCTCTGTGACATCTTTGGAGTCCGGGAGCCATTAAGTCTTCTAGT 815
 DB GGGGACCCCAAGCTCTGTGACATCTTTGGAGTCCGGGAGCCATTAAGTCTTCTAGT 815
 QY 721 CTTATACGTCTCCCATAGGACCAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTG 780
 DB CTTATACGTCTCCCATAGGACCAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTG 875
 QY 816 CTTATACGTCTCCCATAGGACCAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTG 875
 DB CTTATACGTCTCCCATAGGACCAAGGCTTATTCACAAAGCATCATGAGAGTGGGCTG 875
 QY 781 GCCATCATCCCTTACCTGAGGACCATGATTAAGAGAGTGAAGCACTGAGGCTG 840
 DB GCCATCATCCCTTACCTGAGGACCATGATTAAGAGAGTGAAGCACTGAGGCTG 935
 QY 876 GCCATCATCCCTTACCTGAGGACCATGATTAAGAGAGTGAAGCACTGAGGCTG 935
 DB GCCATCATCCCTTACCTGAGGACCATGATTAAGAGAGTGAAGCACTGAGGCTG 935
 QY 841 GCACATTTCTGTGTAAACATGCGTCAAGCTCTGAGGCTCTGAGGCTCTGAGACA 900
 DB GCACATTTCTGTGTAAACATGCGTCAAGCTCTGAGGCTCTGAGGCTCTGAGACA 995
 QY 936 GCACATTTCTGTGTAAACATGCGTCAAGCTCTGAGGCTCTGAGGCTCTGAGACA 995
 DB GCACATTTCTGTGTAAACATGCGTCAAGCTCTGAGGCTCTGAGGCTCTGAGACA 995
 QY 901 AAACCTTCCAGAGACCTGACCTTCAAGGCAAGAAACAAAGTCTTCACTGAGAGTGT 960
 DB AAACCTTCCAGAGACCTGACCTTCAAGGCAAGAAACAAAGTCTTCACTGAGAGTGT 1055
 QY 996 AAACCTTCCAGAGACCTGACCTTCAAGGCAAGAAACAAAGTCTTCACTGAGAGTGT 1055
 DB AAACCTTCCAGAGACCTGACCTTCAAGGCAAGAAACAAAGTCTTCACTGAGAGTGT 1055
 QY 961 GATGTGCTTTCTTCTTAATGAGCCTTAAGATTAATGCTCAGAAAGCATTTAAAGCA 1020

Db 1056 GATGATGCTTCTTCTTCCATAGAGCCTCTAGATATATGTCTCGAAGACATTTAAAGCA 1115
Qy 1021 ATTCCTTCATCATGAGAGTAAATAAACAGAGTGTGCTTCTGCTCCATGAAGAG 1080
Db 1116 ATTCCTTCATCATGAGAGTAAATAAACAGAGTGTGCTTCTGCTCCATGAAGAG 1175
Qy 1081 GCTCCTGAGATCCCTAGTGGCTCCAGAGAGTCCCTGCTCCATGAATGAACATC 1140
Db 1176 GCTCCTGAGATCCCTAGTGGCTCCAGAGAGTCCCTGCTCCATGAATGAACATC 1235
Qy 1141 CTGCACTCCCGCTCAGATATTTGACCTTGTGCTAATGAAATCTTCACAGACAC 1200
Db 1236 CTGCACTCCCGCTCAGATATTTGACCTTGTGCTAATGAAATCTTCACAGACAC 1295
Qy 1201 TCCCTGATGAATCCGAGACAGCTTCTGACCTTGTGCTAATGAAATCTTCACAGAC 1260
Db 1296 TCCCTGATGAATCCGAGACAGCTTCTGACCTTGTGCTAATGAAATCTTCACAGAC 1355
Qy 1261 CTGCACTGATCAAGCTCAGATATCAAGAGATGTGCTGACCTGCTTCTTCTATGAG 1320
Db 1356 CTGCACTGATCAAGCTCAGATATCAAGAGATGTGCTGACCTGCTTCTTCTATGAG 1415
Qy 1321 TTTGGCACCGGCTCAGTCTTTGAAGACAGAGCCGCTTTGTCAAAGCCGACAC 1380
Db 1416 TTTGGCACCGGCTCAGTCTTTGAAGACAGAGCCGCTTTGTCAAAGCCGACAC 1475
Qy 1381 GCTGATGAAGTCCGCTTTGTGCTCGTGGAGCTTCTGAAAGGGGACATTTGATGTC 1440
Db 1476 GCTGATGAAGTCCGCTTTGTGCTCGTGGAGCTTCTGAAAGGGGACATTTGATGTC 1535
Qy 1441 GAAGAGCCACGAGAGAGAGAAATTAAGTGAAGAGATGAAATTAAGTGAAGTGAAC 1500
Db 1536 GAAGAGCCACGAGAGAGAGAAATTAAGTGAAGAGATGAAATTAAGTGAAGTGAAC 1595
Qy 1501 TTTGCTCGAAGCGGGAATCTATGGAAGAGAGCTTCTGCTGCTGCTTATATCTG 1560
Db 1596 TTTGCTCGAAGCGGGAATCTATGGAAGAGAGCTTCTGCTGCTGCTTATATCTG 1655
Qy 1561 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCTGAGACAGAGCTCAAAGACG 1620
Db 1656 ACTGAGCAGTACCTCAGCTGAGCTTGAACATGAGCTGAGACAGAGCTCAAAGACG 1715
Qy 1621 CGGATGATTTTGAACACGACATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 1716 CGGATGATTTTGAACACGACATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 1775
Qy 1681 AGTCTCTTCTTCTTAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1776 AGTCTCTTCTTCTTAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1835
Qy 1741 CTTTGA 1746
Db 1836 CTTTGA 1841

RESULT 3
ABQ86169
ID ABQ86169 standard; DNA; 1728 BP.
AC ABQ86169;
XX 10-SEP-2002 (first entry)
XX
DE Novel human gene. SEQ ID 40.
XX
XX Human; cytosolic; vulnary; antiarteriosclerotic; antiparkinsonian;
KW neurotrophic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiac; antitumor; antitumor; antitumor;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorder; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;

KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyperimmunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo-transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; ss.
XX
OS Homo sapiens.
XX
PN MO200250105-A1.
XX
XX 27-UN-2002.
XX
PF 17-DEC-2001; 2001WO-US049232.
XX
PR 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX (SMK) SMITHKLINE BEECHAM CORP.
PA (SMK) SMITHKLINE BEECHAM PLC.
PA (GLAXO) GLAXO GROUP LTD.
XX
PI Agnew P. Birkeland M, Cogswell JP, Kahnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Serum UC, Xie Q;
XX WPI: 2002-508784/54.
DR P-PSDB; ABP61004.
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
XX Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 2(a); Page 251; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
XX which allow it to be secreted extracellularly or membrane associated. The
XX activity of polypeptides of the invention may be described as,
XX cytosolic, vulnary, antiarteriosclerotic, antiparkinsonian, neurotrophic,
XX neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
XX cardiac, antitumor, antitumor, antitumor, antitumor, antitumor,
XX and metabolic. Polypeptides and polynucleotides of the invention are
XX useful in the treatment, or as a vaccine in the prevention of, cancer,
XX wound healing disorders, infection, atherosclerosis, Parkinson's disease
XX and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
XX inflammation, neoplastic disease, nervous system related disorders and
XX cardiovascular disorders, pancreatitis, respiratory disorder,
XX hyperproliferation, systemic autoimmune disease, hyperimmunity,
XX developmental abnormality, gastrointestinal ulceration, neuropathy,
XX haematological diseases, metabolic diseases, sperm dysfunction, thyroid
XX disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
XX transduction deficiency, neurological diseases, stroke, angiogenesis,
XX trachea, thymus, lymph node and muscular system, obesity, anorexia,
XX growth abnormalities, and alleviation of precocious puberty. The
XX sequences given in records ABQ86130-ABQ86184 represent novel human CDNA's
XX of the invention

SO Sequence 1728 BP; 367 A; 483 C; 437 G; 441 T; 0 U; 0 Other;
XX
XX
Query Match 94.7%; Score 1652.8; DB 6; Length 1728;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1654; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 91 GGAGCTTCTGCTGAAGGGCCACAGAGAAACACAGCTGGATGATTCAGGGCAAGCA 150

```

Db      73 GGGCTTCTGCTGAAAGGCGCAGAGAAACACAGGCTGGATGATTCAGGGCAAGCAA 132
Qy      151 GTCAATGTCGAGGAAAGCCCTGCTGCTGAAACGTGTTCTCTGGAGTCCCTTGTCTGT 210
Db      133 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAAAGTGTCTCTGGAGTCCCTTGTCTGT 192
Qy      211 CCCCCTGAGGATCCCTGCGATTTTACGAACCCGACGCTGATCCGCTGGATTAATTG 270
Db      193 CCCCCCTGGATCCCTGCGATTTTACGAACCCGACGCTGATCCGCTGGATTAATTG 252
Qy      271 CGAAGAGCAGCTCTTACCTTAATTTGTGCTTCAAGACTCAGAGTGGCTGCTTAAGT 330
Db      253 CGAAGAGCAGCTCTTACCTTAATTTGTGCTTCAAGACTCAGAGTGGCTGCTTAAGT 312
Qy      331 CAACATGCTCAAGTGTGATTAACCCGAATTCGGAGTGTCAAGAGATCTGCTTACCTG 390
Db      313 CAACATGCTCAAGTGTGATTAACCCGAATTCGGAGTGTCAAGAGATCTGCTTACCTG 372
Qy      391 AACATCTATGCGCTGCGCCAGCGATACAGGCTCAAGCTCCCGCTTGGTGTGTTTC 450
Db      373 AACATCTATGCGCTGCGCCAGCGATACAGGCTCAAGCTCCCGCTTGGTGTGTTTC 432
Qy      451 CCAGAGGTGCTTCAAGACTGCTCAGCTCAGCTCATCTTGTATGGTTCGCTGCTGCTG 510
Db      433 CCAGAGGTGCTTCAAGACTGCTCAGCTCAGCTCATCTTGTATGGTTCGCTGCTGCTG 492
Qy      511 TATGAGACGTGCTGTGTTGTGCTGCTCAAGTACCGGCTTACGAATTTGTTTCTTCAAC 570
Db      493 TATGAGACGTGCTGTGTTGTGCTGCTCAAGTACCGGCTTACGAATTTGTTTCTTCAAC 552
Qy      571 ACATGGATCAAGCATGCTCCGCGGAACTGGGCTTCAAGAACCAAGTGGCTGCTGCTG 630
Db      553 ACATGGATCAAGCATGCTCCGCGGAACTGGGCTTCAAGAACCAAGTGGCTGCTGCTG 612
Qy      631 TGGGTCCAGAAAGACATGATTTCTTCCGTGGGAAACCCAGCTGTGTGACCATCTTGGC 690
Db      613 TGGGTCCAGAAAGACATGATTTCTTCCGTGGGAAACCCAGCTGTGTGACCATCTTGGC 672
Qy      691 GAGTCCCGGGAGCCATTAAGTGTCTTAAGTCTTAAGTCTTCCATGAGCCAAAGCTTA 750
Db      673 GAGTCCCGGGAGCCATTAAGTGTCTTAAGTCTTAAGTCTTCCATGAGCCAAAGCTTA 732
Qy      751 TTCCCAAGGACCATGAGAGAGTGGGGAGGACCATCATCCCTTACCTGGAGCCCATGAT 810
Db      733 TTCCCAAGGACCATGAGAGAGTGGGGAGGACCATCATCCCTTACCTGGAGCCCATGAT 792
Qy      811 TATGAGAGAGTGAAGACTGAGGAGTGTGACATTTCTGTGTAAACATGCGTCAGAC 870
Db      793 TATGAGAGAGTGAAGACTGAGGAGTGTGACATTTCTGTGTAAACATGCGTCAGAC 852
Qy      871 TCTGAGGCTCTGTGAGGTGCTGAGACCAAAACCTTCCAGAGAGCTGTGACCTCTCAGC 930
Db      853 TCTGAGGCTCTGTGAGGTGCTGAGACCAAAACCTTCCAGAGAGCTGTGACCTCTCAGC 912
Qy      931 CAGAAAACAAAGTCTTTTCACTCGAGTGTGTGATGTGCTTTTCTTAATGAGCTCTTA 990
Db      913 CAGAAAACAAAGTCTTTTCACTCGAGTGTGTGATGTGCTTTTCTTAATGAGCTCTTA 972
Qy      991 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1050
Db      973 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1032
Qy      1051 GAGTGTGCTTCTGTGCTGCTTAATGAAGAGGCTCTGAGATCTTAAGTCTTCAACAG 1110
Db      1033 GAGTGTGCTTCTGTGCTGCTTAATGAAGAGGCTCTGAGATCTTAAGTCTTCAACAG 1092
Qy      1111 TCCCTTGGCTTCCATCTGATACAAACATCCCTGACATCCGCGCTCAAGATTTGACCTT 1170
Db      1093 TCCCTTGGCTTCCATCTGATACAAACATCCCTGACATCCGCGCTCAAGATTTGACCTT 1152
Qy      1171 GTGGCTAAATGATATCTTCATGACAAAGACTCCCTGACTGAATCGAGACAGTCTTCTG 1230

```

```

Db      1153 GTGGCTAAATGATATCTTCATGACAAAGACTCCCTGACTGAATTCAGACAGTCTTCTG 1212
Qy      1231 GACTTGTCTGGAGATGTGTTCTTTGTGCTGCTGACATGATCACAGTCAATATCACAGA 1290
Db      1213 GACTTGTCTGGAGATGTGTTCTTTGTGCTGCTGACATGATCACAGTCAATATCACAGA 1272
Qy      1291 GATGTGTGTGACCTGTCTACTTCTATGAGTTTGGGACCCGCTCAGTGTCTTTGAAGAC 1350
Db      1273 GATGTGTGTGACCTGTCTACTTCTATGAGTTTGGGACCCGCTCAGTGTCTTTGAAGAC 1332
Qy      1351 ACGAAGCCGCTTTTGTGTAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGCTG 1410
Db      1333 ACGAAGCCGCTTTTGTGTAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGCTGCTG 1392
Qy      1411 GCTTCTGAAAGGGGACATTTATGTTTGAAGAGGACGAGAGAGAGAGATTAATCTG 1470
Db      1393 GCTTCTGAAAGGGGACATTTATGTTTGAAGAGGACGAGAGAGAGAGATTAATCTG 1452
Qy      1471 AGCCGGAAGATGATGAATTAATCTGAGGCTACCTTGTCTGAAACCGGAAATCTTAATGGAAC 1530
Db      1453 AGCCGGAAGATGATGAATTAATCTGAGGCTACCTTGTCTGAAACCGGAAATCTTAATGGAAC 1512
Qy      1531 GACTGTCTGTGTGCTCAGCTTATATCTGACTGACGATCACTTCACTGACTGACTTGAAC 1590
Db      1513 GACTGTCTGTGTGCTCAGCTTATATCTGACTGACGATCACTTCACTGACTGACTTGAAC 1572
Qy      1591 ATGAGCTCTGAGACGAGACTCAAGAAACCGGCGGTGATTTTGGACGAGACCATCCGC 1650
Db      1573 ATGAGCTCTGAGACGAGACTCAAGAAACCGGCGGTGATTTTGGACGAGACCATCCGC 1632
Qy      1651 CTGATCTGTGTGCTCTCCAGACATGCTCCAGAGTCTCTTCTTCTTAATCTTCTCTCT 1710
Db      1633 CTGATCTGTGTGCTCTCCAGACATGCTCCAGAGTCTCTTCTTCTTAATCTTCTCTCT 1692
Qy      1711 CTCTCTCAGCTTTCTTTTCTTTTGTGCTCTTGA 1746
Db      1693 CTCTCTCAGCTTTCTTTTCTTTTGTGCTCTTGA 1728

```

```

RESULT 4
AAd40574
ID AAd40574 standard; cDNA; 2232 BP.
XX
XX AAd40574;
AC
XX
XX 30-OCT-2002 (first entry)
DT
XX
XX Human drug metabolizing enzyme (DME-10) cDNA.
DE
XX
XX Human; drug metabolizing enzyme; autoimmune; inflammatory disorder;
XX acquired immunodeficiency syndrome; AIDS; atherosclerosis; psoriasis;
XX proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;
XX asthma; neurological disorder; Alzheimer's disease; Huntington's disease;
XX dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;
XX drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;
XX renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;
XX anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;
XX golitre; gastrointestinal disorder; gene therapy; virulence; liver disorder;
XX anticonvulsant; nootropic; enzyme; DME-10; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1.1929
FT CDS /tag= a
FT /product= "Human DME-10"
FT sig_peptide 1.150
FT mat_peptide 151.1926
FT /tag= b
FT /tag= c
FT /product= "Mature human DME-10"
PN
XX
XX WO200246426-A2.

```

XX 13-JUN-2002.
 PD
 XX
 PF 04-DEC-2001; 2001MO-US047429.
 XX
 PR 08-DEC-2000; 2000US-0254308P.
 XX
 PR 15-DEC-2000; 2000US-0256189P.
 PR 21-DEC-2000; 2000US-0257713P.
 PR 19-JAN-2001; 2001US-0262706P.
 PR 02-FEB-2001; 2001US-0266020P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Sanjanwala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ;
 PI Lee EA, Ding L, Hafalia AJD, Tang YT, Yue H, Tribouley CM, Lu DAM;
 PI Lal PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AK, Lu Y,
 PI Ison CH;
 XX
 DR WPI; 2002-519668/55.
 DR P-PSDB; AAE25025.
 XX
 PT Novel human drug metabolizing polypeptide, useful in diagnosis,
 PT prevention or treatment of autoimmune/inflammatory, cell proliferative,
 PT neurological, developmental, endocrine, metabolic and gastrointestinal
 PT disorders.
 PS
 PS Claim 78; Page 167-168; 169pp; English.
 XX
 XX The invention relates to an isolated human drug metabolizing enzyme (DME)
 CC and its nucleotide. DME is useful for diagnosing, treating or preventing
 CC disorders associated with aberrant expression of DME, where the disorders
 CC are selected from autoimmune/inflammatory disorder such as acquired
 CC immunodeficiency syndrome (AIDS), asthma, atherosclerosis, psoriasis,
 CC uveitis; a cell proliferative disorder such as arteriosclerosis,
 CC cirrhosis, hepatitis, and cancer; a neurological disorder such as
 CC Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;
 CC a developmental disorder such as renal tubular acidosis, epilepsy,
 CC anaemia; an endocrine disorder such as adenoma, thrombosis and infections
 CC ; an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic
 CC disorder such as cystic fibrosis, diabetes and goitre; a gastrointestinal
 CC disorder such as anorexia, peptic ulcer, and liver disorders. DME is
 CC useful in a number of drug screening techniques and to analyse the
 CC proteome of a tissue or cell type. The invention is useful for creating
 CC knock-in humanised animals or transgenic animals to model human diseases,
 CC in somatic or germline gene therapy, to generate a transcript image of a
 CC tissue or cell type, for detecting differences in the chromosomal
 CC location due to translocation, inversion, etc. among normal, carrier or
 CC affected individuals, and as hybridisation probes for mapping naturally
 CC occurring genomic sequences. The present sequence is human DME-10 cDNA
 CC
 XX
 XX Sequence 2232 BP; 504 A; 586 C; 549 G; 593 T; 0 U; 0 Other;
 SQ
 Query Match 94.6%; Score 1651.6; DB 6; Length 2232;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1657; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

DB 504 GCTCTTAGATCAACATGCTCAAGTGCATTACCGGAATTCCGAGTGCAGAAAGACTG
 QY
 DB 381 CCTTACCTGAAACATCTATGCGCCCTGCCACGCCGATACAGGCTCCAGGCTCCCGTCTT
 QY
 DB 564 CTTTACCTGAAACATCTATGCGCCCTGCCACGCCGATACAGGCTCCAGGCTCCCGTCTT
 QY
 DB 441 GGTGTGTTCCAGGAGGTGCTCTTCAAGACTGAGCTCAAGCTCTCAATCTTTGATGGGTCGCG
 QY
 DB 624 GGTGTGTTCCAGGAGGTGCTCTTCAAGACTGAGCTCAAGCTCTCAATCTTTGATGGGTCGCG
 QY
 DB 501 CCTGCTCCCTATGAGACCTGCTGTTGTGTGTGTCAGTACCGGCTAGGAATATTTGG
 DB 684 CTTGCTGCTCGGGTCCAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTCTGTGAC
 QY
 DB 561 TTTCTTACCAACATGAGGATGAGCATGCTCCGGGGAACTGGGCTTCAAGGACCAAGTGGC
 DB 744 TTTCTTACCAACATGAGGATGAGCATGCTCCGGGGAACTGGGCTTCAAGGACCAAGTGGC
 QY
 DB 621 TGCTCTGCTGGGTCCAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTCTGTGAC
 DB 804 TGCTCTGCTGGGTCCAGAAACATGAGTCTTCCGTGGGAGCCCAAGCTCTGTGAC
 QY
 DB 681 CATCTTTGGGAGTCCCGGGAGCCATTAAGTTTCTAAGTTTATCTTCTCCATGGC
 DB 864 CATCTTTGGGAGTCCCGGGAGCCATTAAGTTTCTAAGTTTATCTTCTCCATGGC
 QY
 DB 741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTACCTGGA
 DB 924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTACCTGGA
 QY
 DB 801 GGCCCAATGATTATGAGAAAGTGAAGACTGCAAGTGTGTCACATTTCTGTGTAAACA
 DB 984 GGCCCAATGATTATGAGAAAGTGAAGACTGCAAGTGTGTCACATTTCTGTGTAAACA
 QY
 DB 861 TGCGTCAAGCTTGAAGCCCTGCGTAGGTGCTGAGGACAAACCTCCCAAGAGACTGCT
 DB 1044 TGCGTCAAGCTTGAAGCCCTGCGTAGGTGCTGAGGACAAACCTCCCAAGAGACTGCT
 QY
 DB 921 GACCTTCAGCCAGAAACAAAGTCTTCACTCGAGTGTGATGTGCTTCTTCTTAA
 DB 1104 GACCTTCAGCCAGAAACAAAGTCTTCACTCGAGTGTGATGTGCTTCTTCTTAA
 QY
 DB 981 TGAGCCTTGAATCTATGCTTCAAGAAAGATTTAAAGCAATCTTCCATCATCGAAGT
 DB 1164 TGAGCCTTGAATCTATGCTTCAAGAAAGATTTAAAGCAATCTTCCATCATCGAAGT
 QY
 DB 1041 CAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAGGCTCTGAGATCCTCAGTGG
 DB 1224 CAATTAACAGAGTGTGCTTCTGCTGCTATGAAGAGGCTCTGAGATCCTCAGTGG
 QY
 DB 1101 CTCGACAAAGTCCCTTCCCTCATCTGATTAACAAACATCTTGACATCCGCTCAGTA
 DB 1284 CTCGACAAAGTCCCTTCCCTCATCTGATTAACAAACATCTTGACATCCGCTCAGTA
 QY
 DB 1161 TTGCAACCTTGTGGCTTAATGAATACCTTCATGACAAAGCACTCCGATCGAAATCCGAGA
 DB 1344 TTGCAACCTTGTGGCTTAATGAATACCTTCATGACAAAGCACTCCGATCGAAATCCGAGA
 QY
 DB 1221 CAGTCTTGTGAACCTTGTGAGATGTGTTCTTTGTGTGTCCTGACATGATCAAGCTGG
 DB 1404 CAGTCTTGTGAACCTTGTGAGATGTGTTCTTTGTGTGTCCTGACATGATCAAGCTGG
 QY
 DB 1281 ATATACAGAGATGCTGTGACCTGTCTATCTTATGATTTGGCAACCGGCTCAAGT
 DB 1464 ATATACAGAGATGCTGTGACCTGTCTATCTTATGATTTGGCAACCGGCTCAAGT
 QY
 DB 1341 CTTTGAAGACAGAAAGCCGCTTTTGTCAAGCGACACAGCTGATGAAGTCCGCTTGT
 DB 1524 CTTTGAAGACAGAAAGCCGCTTTTGTCAAGCGACACAGCTGATGAAGTCCGCTTGT
 QY
 DB 1401 GTTCGGTGTGCTTCTTGAAGGGGACATTTGATGTTGAAAGGACACGAGAGAGA 1460


```

QY 921 GACCTGAGCAGAAACAAAGTCTTCACTGAGTGTGATGAGTCTTCTTCTTAA 980
DB 1103 GACCTGAGCAGAAACAAAGTCTTCACTGAGTGTGATGAGTCTTCTTCTTAA 1162
QY 981 TGAGCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1040
DB 1163 TGAGCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1222
QY 1041 CAATTAACAGAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
DB 1223 CAATTAACAGAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1282
QY 1101 CTCGCAAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160
DB 1283 CTCGCAAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1342
QY 1161 TTGTCACCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTG 1220
DB 1343 TTGTCACCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTG 1402
QY 1221 CAGTCTTCTGAGCTTCTGAGAGTGTGCTTGTGTGCTGCTGCTGCTGCTGCTG 1280
DB 1403 CAGTCTTCTGAGCTTCTGAGAGTGTGCTTGTGTGCTGCTGCTGCTGCTGCTG 1462
QY 1281 ATATCAAGAGATGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1340
DB 1463 ATATCAAGAGATGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1522
QY 1341 CTTTGAAGACAGAGACCGGCTTTGTCGAGGCGGACCGCTGATGAGTCCGCTTGT 1400
DB 1523 CTTTGAAGACAGAGACCGGCTTTGTCGAGGCGGACCGCTGATGAGTCCGCTTGT 1582
QY 1401 GTTCGCTGCTGCTTCTGTCGAGGCGGACATGTTGTTGTCGAGGCGGACCGGAG 1460
DB 1583 GTTCGCTGCTGCTTCTGTCGAGGCGGACATGTTGTTGTCGAGGCGGACCGGAG 1642
QY 1461 GAGTTCACGAGCCGAGAGATGATGAAATCTGCGCTACCTTTGCTCGAACCGGAA 1520
DB 1643 GAGTTCACGAGCCGAGAGATGATGAAATCTGCGCTACCTTTGCTCGAACCGGAA 1702
QY 1521 TAATGGAACGAGCTGTCTGTGTCGAGCTTATATGTCGAGTGAAGTACCTCACT 1580
DB 1703 TAATGGAACGAGCTGTCTGTGTCGAGCTTATATGTCGAGTGAAGTACCTCACT 1762
QY 1581 GAGCTTGAACATGAGCTTCTGTCGAGAGACTCAAGAACCGCGGTGATTTTGGAC 1640
DB 1763 GAGCTTGAACATGAGCTTCTGTCGAGAGACTCAAGAACCGCGGTGATTTTGGAC 1822
QY 1641 CACCATCCCGGATCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1700
DB 1823 CACCATCCCGGATCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1882
QY 1701 TTTCCTCTCTCTCTCTGAGCTTTCTTTCTTTGTCGCTCTGGA 1746
DB 1883 TTTCCTCTCTCTCTCTGAGCTTTCTTTCTTTGTCGCTCTGGA 1928

```

RESULT 6
ADB62095
ID ADB62095 standard; cDNA; 2092 BP.

ADB62095;

04-DEC-2003 (first entry)

Human cDNA encoding clone BRAWH20021910.

Human; 88; gene; pharmaceutical; diagnostic; gene therapy;
tissue regeneration; cell regeneration; membrane protein;
signal transduction-related protein; transcription-related protein;
osteoporosis; neurological disease; cancer; tumour.

OS Homo sapiens.

Key Location/Qualifiers
CDS 394..1803

FT /tag= a
FT /product= "Clone BRAWH20021910 protein"

PN BP1308459-A2.

XX 07-MAY-2003.

PF 28-MAR-2002; 2002EP-00007401.

PR 05-NOV-2001; 2001JP-00379298.

XX 25-JAN-2002; 2002US-00350978.

PA (HELI-) HELIX RES INST.
(REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y,

XX WPI; 2003-450961/43.

DR P-PSDB; ADB64065.

PT New polynucleotides and polypeptides, useful for developing a diagnostic
marker or medicines for regulation of their expression and activity, or
as targets of gene therapy.

PS Claim 1; Page; 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
from 1970 fully defined nucleotide sequences which encode novel
polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
or its partial peptide, an antibody binding to the polypeptide or peptide
of the polynucleotide, immunologically assaying the polypeptide or
peptide of the polynucleotide by contacting the polypeptide or peptide
with the antibody of the encoded protein, and observing the binding
between the two, a transformant carrying the polynucleotide in an
expressible manner and an antisense polynucleotide. The oligonucleotide
is useful as a primer for synthesizing the polynucleotide, or as a probe
for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transduction-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.
XX

Sequence 2092 BP; 464 A; 552 C; 497 G; 579 T; 0 U; 0 Other;

Query Match 93.4%; Score 1631.2; DB 10; Length 2092;

Best Local Similarity 99.2%; Pred. No. 0;
Matches 1653; Conservative 0; Mismatches 3; Indels 10; Gaps 1;

QY 91 GGGCTTCTGTCGAGAGGCGCAGAGAGACACAGGCTGGATGATGAGGAGCA 150

DB 138 GGGCTTCTGTCGAGAGGCGCAGAGAGACACAGGCTGGATGATGAGGAGCA 197

QY 151 GTCACTGTCTGGAGAGCCCTGTGCTGGAACGTGTTCTTGGAGTCCCTTTGCTGT 210

DB 198 GTCACTGTCTGGAGAGCCCTGTGCTGGAACGTGTTCTTGGAGTCCCTTTGCTGT 257

QY 211 CCCCCGCTGGAGTCCCTGTGATTTAGAACCCGAGCTGCAATGCGCTTGGATTA 270

DB 258 CCCCCGCTGGAGTCCCTGTGATTTAGAACCCGAGCTGCAATGCGCTTGGATTA 317

271 CGAAGGACCTCTTACCTTAATTT-----GTGCTCCAGAACTCAGAGTGGCT 320
318 CGAAGGACCTCTTACCTTAATTTTGAAGACAGAGTGGCTCCAGAACTCAGAGTGGCT 377
321 GCTCTTAAGTAAACAATGCTCAAGTGTGATTAACCGAATTCGAGTGTCAAGAACTG 380
378 GCTCTTAAGTAAACAATGCTCAAGTGTGATTAACCGAATTCGAGTGTCAAGAACTG 437
381 CCTCTAAGTAAACAATGCTCAAGTGTGATTAACCGAATTCGAGTGTCAAGAACTG 440
438 CCTCTAAGTAAACAATGCTCAAGTGTGATTAACCGAATTCGAGTGTCAAGAACTG 497
441 GGTGTGTGTCCAGAGAGGCTCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 500
498 GGTGTGTGTCCAGAGAGGCTCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 557
501 CCTGTGTGTCCAGAGAGGCTCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 560
558 CCTGTGTGTCCAGAGAGGCTCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 617
561 TTTCTTCAACACATGAGATCAGCATGCTCCGAGGAACTGAGCTTCAAGAACAGTGGC 620
618 TTTCTTCAACACATGAGATCAGCATGCTCCGAGGAACTGAGCTTCAAGAACAGTGGC 677
621 TGTCTGTGTCCAGAGAGGCTCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 680
678 TGTCTGTGTCCAGAGAGGCTCTCAAGCTCAAGCTCAAGCTCAAGCTCAAGCTCAAG 737
681 CATCTTTGGGAGATCCGGGAGGAGCATTAAGTCTTCAAGTCTTCAAGTCTTCAAGTGC 740
738 CATCTTTGGGAGATCCGGGAGGAGCATTAAGTCTTCAAGTCTTCAAGTCTTCAAGTGC 797
741 CAAGGCTTTATTCACAAAGCATCAATGAGAGTGGAGTGGCCATCATCTTACCTGGA 800
798 CAAGGCTTTATTCACAAAGCATCAATGAGAGTGGAGTGGCCATCATCTTACCTGGA 857
801 GGGCCATATTAATGAGAGAGTGAAGACCTGCAAGTGTGTGACAAATTCCTGTGTGAACA 860
858 GGGCCATATTAATGAGAGAGTGAAGACCTGCAAGTGTGTGACAAATTCCTGTGTGAACA 917
861 TGGCTCAGACTCTGAGGCTCTGCTGAGTGTGCTGAGAGCAAAACCTCCAGAGAGTGT 920
918 TGGCTCAGACTCTGAGGCTCTGCTGAGTGTGCTGAGAGCAAAACCTCCAGAGAGTGT 977
921 GACCTTCAAGCAGAAACAAAGTCTTCACTCAAGTGTGAGTGTCTTCTTCTTAA 980
978 GACCTTCAAGCAGAAACAAAGTCTTCACTCAAGTGTGAGTGTCTTCTTCTTAA 1037
981 TGAAGCTCTAATCTAATGTCTCAAGAAAGATTTTAAAGCAATTCCTTCATCATCGAGAT 1040
1038 TGAAGCTCTAATCTAATGTCTCAAGAAAGATTTTAAAGCAATTCCTTCATCATCGAGAT 1097
1041 CAATTAACAGAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1100
1098 CAATTAACAGAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157
1101 CTCCAAAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1160
1158 CTCCAAAGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1217
1161 TTTTGACCTTGTGTGCTAATGATATCTTCATGACAAAGCACTCCCTGACGAAATCCGAGA 1220
1218 TTTTGACCTTGTGTGCTAATGATATCTTCATGACAAAGCACTCCCTGACGAAATCCGAGA 1277
1221 CAGTCTTGTGACCTTGTGCTGAGATGTGTCTTGTGTGCTGCTGCTGCTGCTGCTGCTG 1280
1278 CAGTCTTGTGACCTTGTGCTGAGATGTGTCTTGTGTGCTGCTGCTGCTGCTGCTGCTG 1337
1281 ATATCAAGAGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1340
1338 ATATCAAGAGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1397

1341 CTTGAAGACAGAAAGCCGCTTTTGTCAAAACCGACAGCTGATGAAGTCCGCTTGT 1400
1398 CTTGAAGACAGAAAGCCGCTTTTGTCAAAACCGACAGCTGATGAAGTCCGCTTGT 1457
1401 GTTCGCTGTGCTCTTCTGAAAGGAGCAATTTTATGTTTGAAGAGCAAGAGAGGA 1460
1458 GTTCGCTGTGCTCTTCTGAAAGGAGCAATTTTATGTTTGAAGAGCAAGAGAGGA 1517
1461 GAAATTACTGAGCCGAGAGATGATGAATATCTGGGCTACCTTGTCTGAAACGGGAAATC 1520
1518 GAAATTACTGAGCCGAGAGATGATGAATATCTGGGCTACCTTGTCTGAAACGGGAAATC 1577
1521 TAATGGAAAGCACTGTCTGTGAGCACTTATATCTGATGAGCACTGACCTTCAAGCT 1580
1578 TAATGGAAAGCACTGTCTGTGAGCACTTATATCTGATGAGCACTGACCTTCAAGCT 1637
1581 GGAATTGAACATGAGCTCTGAGCAAGACTCAAGAACCGCGGAGATTTTGAACAG 1640
1638 GGAATTGAACATGAGCTCTGAGCAAGACTCAAGAACCGCGGAGATTTTGAACAG 1697
1641 CACCATCCCTGATATCTGTGCTGCTCCGACATGCTCCACAGTCTTCTTCTTAC 1700
1698 CACCATCCCTGATATCTGTGCTGCTCCGACATGCTCCACAGTCTTCTTCTTAC 1757
1701 TTTCTCTCTCTCTCTCAGACCTTCTTCTTCTTGTGTGCTCTTGA 1746
1758 TTTCTCTCTCTCTCTCAGACCTTCTTCTTCTTGTGTGCTCTTGA 1803

RESULT 7
ABX72267
ID ABX72267 strand: cDNA, 1746 BP.
XX
AC ABX72267;
XX
XX 03-JUN-2003 (first entry)
DT
XX
XX
DE Human NOVX polynucleotide #98.
XX
KW Human; NOVX; gene; ss; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; arrhythmogenic right ventricular dysplasia; ASD;
KW pulmonary stenosis; subcoronary stenosis; ventricular septal defect; VSD;
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
OS Homo sapiens.
XX
XX W0200281498-A2.
PN
XX
PD 17-OCT-2002.
PF
PF 03-APR-2002; 2002MO-US010780.
XX
XX 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282020P.
PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 13-APR-2001; 2001US-0284234P.
PR 17-APR-2001; 2001US-0285325P.
PR 19-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.

PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324802P.
PR 25-SEP-2001; 2001US-0324802P.
PR 17-OCT-2001; 2001US-0325684P.
PR 14-NOV-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 21-NOV-2001; 2001US-0332779P.
PR 04-DEC-2001; 2001US-0332115P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
(CURA-) CURAGEN CORP.

PI Guo X, Kekuda R, Miller CE, Malyankar UM, Szytek KA;
PI Paturajan M, Liu X, Gusev VY, Li Y, Vernet CM, Zernhusen BD;
PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
PI Padigar M, Shimbets RA, Gangoli EA, Taupier RJ, Caeman SJ, Ji W;
PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ;
PI MacDougall JR, Rothenberg ME, Mazur A, Miller I, Peyman JA;
PI Ellerman K;

DR MPI: 2003-046858/04.
DR P-PSDB; AB054639.

PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.

PS Claim 17; Page 303; 666pp; English.

XX The invention relates to human polypeptides, termed NOVX, and the
XX polynucleotides encoding them. The polypeptides and polynucleotides are
XX useful for diagnosing disease, and screening for potential therapeutic
XX agents. The sequences are useful for treating metabolic disorders,
XX cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
XX stenosis, atrial septal defect (ASD), atrioventricular canal defect,
XX ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
XX septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
XX atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
XX disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
XX haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
XX and cancer. Sequences ABX72170-ABX72275 represent human NOVX
XX polynucleotides of the invention

XX Sequence 1746 BP; 402 A; 457 C; 459 G; 428 T; 0 U; 0 Other;

Query Match 82.2%; Score 1435.6; DB 8; Length 1746;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 59; Indels 15; Gaps 1;

QY 81 GGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAACACAGGGCTGGATGATTTCA 140
DB 69 GGTCAAGAGAGGGGCTTCTGCTGAAGGGCCACAGAGAACACAGGGCTGGATGATTTCA 128
QY 141 GGGCAAGACAGTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCCTCGAGTCCC 200
DB 129 GGGCAACAGTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCCTCGAGTCCC 188
QY 201 CTTTGCTGCTCCCGGCTGGAGATCCCTGCGATTACGAAACCCGAGCTGTGATGCTG 260
DB 189 CTTTGCTGCTCCCGGCTGGAGATCCCTGCGATTACGAAACCCGAGCTGTGATGCTG 248
QY 261 GGATTAACCTTGAGAGACCACTTCTAACCCTAATTGTGTGCTTCAGAACTCAGAGTGGCT 320

DB 249 GGATTAACCTTGAGAGACCACTTCTAACCCTAATTGTGTGCTTCAGAACTCAGAGTGGCT 308
QY 321 GCTTTAGATCAACATGCTCAAGTGCATTAACCCGAATTCGAGTGTGAGAGACTG 380
DB 309 GCTTTAGATCAACATGCTCAAGTGCATTAACCCGAATTCGAGTGTGAGAGACTG 368
QY 381 CTTTAACCTGAACATGATGCGCTTGGCCCAAGCGATACAGAGCTTCAAGCTCCCGCTT 440
DB 369 CTTTAACCTGAACATGATGCGCTTGGCCCAAGCGATACAGAGCTTCAAGCTCCCGCTT 428
QY 441 GGTGTGTTCCAGAGAGGCTTCAAGAGCTGAGCTCAAGCTTCAAGTGTGCTCCG 500
DB 429 GGTGTGTTCCAGAGAGGCTTCAAGAGCTGAGCTCAAGCTTCAAGTGTGCTCCG 488
QY 501 CTTGCTGCTCAATAGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 560
DB 489 CTTGCTGCTCAATAGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 548
QY 561 TTTCTTCAACATAGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 620
DB 549 TTTCTTCAACATAGAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 608
QY 621 TGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 680
DB 609 TGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 668
QY 681 CATCTTTGGCGAGTCCCGGGAGCCATAGTGTGTTTCACTTCACTTCACTTCACTTCACTT 740
DB 669 CATCTTTGGCGAGTCCCGGGAGCCATAGTGTGTTTCACTTCACTTCACTTCACTTCACTT 728
QY 741 CAAAGGCTTATTCACAAAGCCATCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 800
DB 729 CAAAGGCTTATTCACAAAGCCATCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 788
QY 801 GGGCCATGATTATGAGAGAGTGAAGAGCTGAGAGTGTGACATTTCTGTGTGAACA 860
DB 789 GGGCCATGATTATGAGAGAGTGAAGAGCTGAGAGTGTGACATTTCTGTGTGAACA 848
QY 861 TGGCTGAGACTTGAAGGCTTGTGAGTGTGCTGAGAGCAAAACCTTCAAGAGAGTGT 920
DB 849 TGGCTGAGACTTGAAGGCTTGTGAGTGTGCTGAGAGCAAAACCTTCAAGAGAGTGT 908
QY 921 GACCTCAGCAGAAAACAAGCTTTTCACTGAGTGTGATGAGTGTGCTTCTTCTTA 980
DB 909 GACCTCAGCAGAAAACAAGCTTTTCACTGAGTGTGATGAGTGTGCTTCTTCTTA 968
QY 981 TGAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040
DB 969 TGAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028
QY 1041 CAATTAACACAGAGTGTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1100
DB 1029 CAATTAACACAGAGTGTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1088
QY 1101 CTCACAAAGTCCCTGCTCCATCTGATACAA-----ACATCTGCA 1145
DB 1089 CACTCCCTCAACCGATGAGCACTTGGCTTCAACAGCTGGGCACTTCCACAGAGACA 1148
QY 1146 CATCCGGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205
DB 1149 TATCCGGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
QY 1206 GACTGAATCCAGACAGTCTTCTGAGCTTGTGAGATGATGATGATGATGATGATGATGATGAT 1265
DB 1209 GACTGAATCCAGACAGTCTTCTGAGCTTGTGAGATGATGATGATGATGATGATGATGATGAT 1268
QY 1266 ACTGATCAAGCTCAATATCAACAGAGTCTGATGACCTGTCTACTTATGATGATGATGATGAT 1325
DB 1269 ACTGATCAAGCTCAATATCAACAGAGTCTGATGACCTGTCTACTTATGATGATGATGATGAT 1328
QY 1326 GACACGGGCTCAAGTGTGATGAGACAGAGGCGCTTTGTCAAGCCGACAGAGCTGA 1385

Db 1329 GCACCGGCTCAGTCTTTGAGACAGACAGCCAGCTTTTGTCAAGCCGACGCTGA 1388
Qy 1386 TGAAGTCCGCTTTGTTGTTGCGTGGTCCCTTCTGAAAGGGGGA CATTTGTAATTTGAAAG 1445
Db 1389 TGAAGTCCGCTTTGTTGTTGCGTGGTCCCTTCTGAAAGGGGGA CATTTGTAATTTGAAAG 1448
Qy 1446 AGCCACGAGAGAGAGAAATTTACTGAGCCGGAAGATGATGAATATCTGGGCTTACCTTTC 1505
Db 1449 AGCCACGAGAGAGAGAAATTTACTGAGCCGGAAGATGATGAATATCTGGGCTTACCTTTC 1508
Qy 1506 TCGAATCCGGAATCTTAATGGAGACGACTGTCTGTGGCCAGCTTAATTTGATCTGA 1565
Db 1509 TCGAATCCGGAATCTTAATGGAGACGACTGTCTGTGGCCAGCTTAATTTGATCTGA 1568
Qy 1566 GCAGTACTCCAGCTGGAATTGAATAGAGCTTGGACAGAGACTCAAGAAACCCCGGT 1625
Db 1569 GCAGTACTCCAGCTGGAATTGAATAGAGCTTGGACAGAGACTCAAGAAACCCCGAG 1628
Qy 1626 GGAATTTTGG 1635
Db 1629 AGATGTGTGG 1638

RESULT 8
ABQ86170 ID ABQ86170 standard; DNA; 1746 BP.
AC ABQ86170;
XX 10-SEP-2002 (first entry)
DT XX
DE Novel human gene. SEQ ID 41.

Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian;
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiac; antidiabetic; antihypertensive;
KW cerebroprotective; anorectic; metabolic; vaccinal; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disease; nervous system disorder;
KW inflammation; neoplastic disease; hyperproliferation; hyperimmunity;
KW hyperproliferation; systemic autoimmune disease; hyperinflammation;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW come photo-transduction deficiency; neurological disease; stroke;
KW angiodysplasia; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty; gene; 88.

OS Homo sapiens.
XX
XX MO200250105-A1.
XX
XX 27-JUN-2002.
XX
XX 17-DEC-2001; 2001WO-US049232.
XX
XX 19-DEC-2000; 2000US-0256710P.
XX 20-DEC-2000; 2000US-0257048P.
XX 09-JAN-2001; 2001US-0260482P.
XX 30-JAN-2001; 2001US-0264922P.
XX 06-FEB-2001; 2001US-0266797P.
XX 19-MAR-2001; 2001US-0276988P.
XX 04-APR-2001; 2001US-0281535P.
XX 08-MAY-2001; 2001US-0289622P.
XX
XX (SMK) SMITHKLINE BEECHAM CORP.
XX (SMK) SMITHKLINE BEECHAM PLC.
XX (GLAXO) GLAXO GROUP LTD.
XX
XX Agawal P, Birkeland M, Cogswell JP, Kabinick KF, Lai Y,
XX Martensen SA, Ritzvi SK, Smith RF, Strum JC, Xie Q,

XX WPI: 2002-508784/54.
DR P-PSDB; ABP61005.
XX
XX Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 2(a); Page 251-252; 335pp; English.

XX The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytostatic, vulnary, antiarteriosclerotic, antiparkinsonian, neurotropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiac, antidiabetic, antihypertensive, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder.
CC hyperproliferation, systemic autoimmune disease, hyperimmunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, come photo-
CC transduction deficiency, neurological diseases, stroke, angiodysplasia,
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
CC of the invention

XX
XX Sequence 1746 BP; 401 A; 464 C; 459 G; 422 T; 0 U; 0 Other;
SQ

Query Match 82.0%; Score 1432.4; DB 6; Length 1746;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 51; Indels 21; Gaps 1;

Qy 91 GGGCTTCTGCTGAAAGGGCCACAGAGAAACACAGAGCTGGATTCAGGGCAAGCA 150
Db 73 GGGCTTCTGCTGAAAGGGCCACAGAGAAACACAGAGCTGGATTCAGGGCAAGCA 132
Qy 151 GTCACTGTGCTGGAAAGCCCTGTGCTGAAAGCTGCTGCTGCTGCTGCTGCT 210
Db 133 GTCACTGTGCTGGAAAGCCCTGTGCTGAAAGCTGCTGCTGCTGCTGCTGCT 192
Qy 211 CCCCCGTGGAGTCCCTGCGATTTACGAACCCGACGCTGATCCCTGGATTAATTG 270
Db 193 CCCCCGTGGAGTCCCTGCGATTTACGAACCCGACGCTGATCCCTGGATTAATTG 252
Qy 271 CGAAGAGCCACTCTTACCTTAATTTGCTCCAGAACTCAGAGTGGCTCTTAAT 330
Db 253 CGAAGAGCCACTCTTACCTTAATTTGCTCCAGAACTCAGAGTGGCTCTTAAT 312
Qy 331 CAACATGCTCAAGGTGATTAACCGGAATTCGAGGTGCGAAGACTGCTTACCTG 390
Db 313 CAACATGCTCAAGGTGATTAACCGGAATTCGAGGTGCGAAGACTGCTTACCTG 372
Qy 391 AACATCTATGCGCTGCCACGCGATACAGGCTCAAGCTCCCGCTTGTGTGCTTC 450
Db 373 AACATCTATGCGCTGCCACGCGATACAGGCTCAAGCTCCCGCTTGTGTGCTTC 432
Qy 451 CCAGAGGTGCTTCAAGACTGCTCAGCTCAGCTTCTTATGAGTCCGCTGCTGCTC 510
Db 433 CCAGAGGTGCTTCAAGACTGCTCAGCTCAGCTTCTTATGAGTCCGCTGCTGCTC 492
Qy 511 TATGAGAGCGTCTGT 570
Db 493 TATGAGAGCGTCTGT 552
Qy 571 ACATGGATTCAGCATCTCCGGGAACTGGGCTTCAAGAGCAGAGTGGCTGCTGTGCT 630

Query Match 79.4%; Score 1386.2; DB 7; Length 1857;
 Best Local Similarity 94.7%; Pred. No. 0;
 Matches 1473; Conservative 0; Mismatches 13; Indels 69; Gaps 1;

```

QY 81 GGGAAAACTGGGCTTCTGTGTAAGGCGCAGAGAGAACACAGGCTGGATGATTTCA 140
DB 264 GGTCCAAAGAGGGCTTCTGTGTAAGGCGCAGAGAGAACACAGGCTGGATGATTTCA 323
QY 141 GGGCAAGCAAGTCACTGTGTGGGAAAGCCCTGTGCTTGAACGTGTTCCTGGAGTCCC 200
DB 324 GGGCAAGCAAGTCACTGTGTGGGAAAGCCCTGTGCTTGAACGTGTTCCTGGAGTCCC 383
QY 201 CTTTGTGTCCCGCTGGAGATCCCTGGAGTTTAAAGAACCGGACCTGCAGATCGCCCTG 260
DB 384 CTTTGTGTCCCGCTGGAGATCCCTGGAGTTTAAAGAACCGGACCTGCAGATCGCCCTG 443
QY 261 GGATAACTTGGAGAGCACTCTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCT 320
DB 444 GGATAACTTGGAGAGCACTCTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCT 503
QY 321 GCTTTAGATCAACATGCTCAAGTGCAATTAACCGAAATTCGAGGTCAAGAGACTG 380
DB 504 GCTTTAGATCAACATGCTCAAGTGCAATTAACCGAAATTCGAGGTCAAGAGACTG 563
QY 381 CCTTACCTGAACATCTATGCGCTGGCCCAACCGGATAGAGGCTCCAGCTCCGCTTT 440
DB 564 CCTTACCTGAACATCTATGCGCTGGCCCAACCGGATAGAGGCTCCAGCTCCGCTTT 623
QY 441 GGTGTGTGTTCCAGAGAGTGTCTTCAAGACTGAGCTCAGCTTCATTTGATGGTCCG 500
DB 624 GGTGTGTGTTCCAGAGAGTGTCTTCAAGACTGAGCTCAGCTTCATTTGATGGTCCG 683
QY 501 CTTGCTGTCTATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 560
DB 684 CTTGCTGTCTATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 743
QY 561 TTTTCTTACCAATGAGGATCAGACATGCTCCGGGAGAACTGGGCTTCAAGAGCAAGTGGC 620
DB 744 TTTTCTTACCAATGAGGATCAGACATGCTCCGGGAGAACTGGGCTTCAAGAGCAAGTGGC 803
QY 621 TGTCTGTGCTGGGTTCAGAGAAACATCGAGTTCCTGAGGAGAACCCAGCTGTGTGAC 680
DB 804 TGTCTGTGCTGGGTTCAGAGAAACATCGAGTTCCTGAGGAGAACCCAGCTGTGTGAC 863
QY 681 CATCTTTGGCGAGTCCGCGGAGCAATAAGTGTTCATGCTTAACTGTCTCCATGGC 740
DB 864 CATCTTTGGCGAGTCCGCGGAGCAATAAGTGTTCATGCTTAACTGTCTCCATGGC 923
QY 741 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTAAGTGA 800
DB 924 CAAAGGCTTATTCACAAAGCCATCATGAGAGTGGGGTGGCCATATCCCTTAAGTGA 983
QY 801 GGCCCATGATTAATGAGAGAGTGAAGACTGTGAGAGGAGTGTGACATTTCTGTGTAAACA 860
DB 984 GGCCCATGATTAATGAGAGAGTGAAGACTGTGAGAGGAGTGTGACATTTCTGTGTAAACA 1043
QY 861 TGGGTCAAGCTGTGAGGCGCTGTGAGGTGCTGTGAGAGCAAAACCTTCAGAGAGTGTCT 920
DB 1044 TGGGTCAAGCTGTGAGGCGCTGTGAGGTGCTGTGAGAGCAAAACCTTCAGAGAGTGTCT 1103
QY 921 GACCTCAGCCGAGAAACAAAGTCTTTCACTGAGTGTGTGATGTGTGTCTTTCTTAA 980
DB 1104 GACCTCAGCCGAGAAACAAAGTCTTTCACTGAGTGTGTGATGTGTGTCTTTCTTAA 1163
QY 981 TGAGGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGAT 1040
DB 1164 TGAGGCTCTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGAT 1223
QY 1041 CAATAACCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1100
DB 1224 CAATAACCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1255

```

```

QY 1101 CTCCAACAAGTCCCTTGTCCCTCCATCTGATATACAAAAATCTGTGACATCCGCTCAGTA 1160
DB 1256 -----TGCATATCCGCTCAGTA 1274
QY 1161 TTTGCACTTGTGGCTTAATGAATPACTTCATGACAAAGCACTCCCTGACTGAATCCGAGA 1220
DB 1275 TTTGCACTTGTGGCTTAATGAATPACTTCATGACAAAGCACTCCCTGACTGAATCCGAGA 1334
QY 1221 CAGTCTTGTGACTTGTGAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1280
DB 1335 CAGTCTTGTGACTTGTGAGATGTGTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1394
QY 1281 ATATCACAAGAGATGTGTGACCTGTCTACTTCTATAGATTTCGACACCGGCTCAGTG 1340
DB 1395 ATATCACAAGAGATGTGTGACCTGTCTACTTCTATAGATTTCGACACCGGCTCAGTG 1454
QY 1341 CTTTGAAAGCAGAAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1400
DB 1455 CTTTGAAAGCAGAAAGCCGCTTTTGTCAAGCCGACACGCTGATGAAGTCCGCTTGT 1514
QY 1401 GTTGGGTGTGCTTCTGTAAGGGGACATTTGTATGTTTCAAGAGACCAAGAGAGAGA 1460
DB 1515 GTTGGGTGTGCTTCTGTAAGGGGACATTTGTATGTTTCAAGAGACCAAGAGAGAGA 1574
QY 1461 GAAGTACTGAGCCGGAAGATGATGAATACTGGGCTACCTTGTCTGAAACCGGAAATCC 1520
DB 1575 GAAGTACTGAGCCGGAAGATGATGAATACTGGGCTACCTTGTCTGAAACCGGAAATCC 1634
QY 1521 TAATGGGAACGACCTGTCTCTGTGGCCAGCTTATAATCTGACTGAGACGACTTCAGCT 1580
DB 1635 TAATGGGAACGACCTGTCTCTGTGGCCAGCTTATAATCTGACTGAGACGACTTCAGCT 1694
QY 1581 GGACTTGAACATGAGCCTGTGACAGAGACTCAAGAAACCGCGGGTGAATTTTGG 1635
DB 1695 GGACTTGAACATGAGCCTGTGACAGAGACTCAAGAAACCGCGGGTGAATTTTGG 1749

```

RESULT 10
 ADA53168
 ID ADA53168 standard; cDNA; 1962 BP.
 XX
 AC ADA53168;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 XX
 DE Human coding sequence, SEQ ID 736.
 XX
 XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW Inflammatory disease; osteoporosis; neurological disease; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 XX 21-MAR-2002; 2002EP-00006586.
 PF 14-SBP-2001; 2001JP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
 XX
 DR WPI; 2003-395539/38.
 DR P-PSDB; ADA54807.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory

PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 PS Claim 1; SEQ ID NO 736; 205bp; English.

CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA5433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 1962 BP; 428 A; 541 C; 473 G; 520 T; 0 U; 0 Other;

Query Match 77.0%; Score 1344.4; DB 10; Length 1962;
 Best Local Similarity 90.9%; Pred. No. 0;
 Matches 1505; Conservative 0; Mismatches 1; Indels 150; Gaps 1;

```

QY 91 GGGCCCTTCCTGTAAGAGCCACAGAGAAACACAGGCTGGATGATTACAGGCAAGCAA 150
DB 222 GGGCCCTTCCTGTAAGAGCCACAGAGAAACACAGGCTGGATGATTACAGGCAAGCAA 281
QY 151 GTCACTGTCTGGGAAGCCCTGTGCTGTGAACGTCTTCTGGAGTCCCTTTGCTGCT 210
DB 282 GTCACTGTCTGGGAAGCCCTGTGCTGTGAACGTCTTCTGGAGTCCCTTTGCTGCT 341
QY 211 CCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATCGCCCTGGATTAATTG 270
DB 342 CCCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATCGCCCTGGATTAATTG 401
QY 271 CGAAGACCAACCTCCACCTTAATTGTCCTCCAGAACTCAGAGTGGTCTCTTAAT 330
DB 402 CGAAGACCAACCTCCACCTTAATTGTCCTCCAGAACTCAGAGTGGTCTCTTAAT 461
QY 331 CAACA CATGCTCAAGGTGATTACCGAAATTCGAGTGTCAAGAACTGCTCTACCG 390
DB 462 CAACA CATGCTCAAGGTGATTACCGAAATTCGAGTGTCAAGAACTGCTCTACCG 521
QY 391 AACATCTATGCGCTGCCCAAGCCGATACAGGCTCCAGACTCCCGTCTGGTGTTC 450
DB 522 AACATCTATGCGCTGCCCAAGCCGATACAGGCTCCAGACTCCCGTCTGGTGTTC 581
QY 451 CCAGAGGTGCTCTTAAGACTGCTCAGGCTCATCTTGAATGGGTCGGCCCTGCTCC 510
DB 582 CCAGAGGTGCTCTTAAGACTGCTCAGGCTCATCTTGAATGGGTCGGCCCTGCTCC 641
QY 511 TATGAGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 570
DB 642 TATGAGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 701
QY 571 ACATGGGATCAGCATGCTCCGGGGAATCGGCTTCAAGACCAAGTGGCTGTGTGT 630
DB 702 ACATGGGATCAGCATGCTCCGGGGAATCGGCTTCAAGACCAAGTGGCTGTGTGT 761
QY 631 TGGGTCCAGAAAGACATCAGATCTTCGATGGGGAACCCAGCTGTGTAACCATTTGGC 690
DB 762 TGGGTCCAGAAAGACATCAGATCTTCGATGGGGAACCCAGCTGTGTAACCATTTGGC 821
QY 691 GAGTCCCGGGAGCCATAGTGTTCATGCTTATATCTGTCTCCATGGCCAAAGCTTA 750
DB 822 GAGTCCCGGGAGCCATAGTGTTCATGCTTATATCTGTCTCCATGGCCAAAGCTTA 881
QY 751 TTCCACAAGCCATCATGAGAGTGGGTGGCATCATCCCTTACCTGAGAGCCCATGAT 810
DB 882 TTCCACAAGCCATCATGAGAGTGGGTGGCATCATCCCTTACCTGAGAGCCCATGAT 941
QY 811 TATGAGAAGATGAGACCTGCAAGTGGTGTGCAATTTCTGTGTAACTATGCTCAGAC 870
DB 942 TATGAGAAGATGAGACCTGCAAGTGGTGTGCAATTTCTGTGTAACTATGCTCAGAC 1001
QY 871 TCTGAGGCCCTGCTGAGAGTGTGCTGAGACAAACCTTCAGAGAGCTGTGACCTCAGC 930
DB 1002 TCTGAGGCCCTGCTGAGAGTGTGCTGAGACAAACCTTCAGAGAGCTGTGACCTCAGC 1061

```

```

QY 931 CAGAAAACAAAGCTTTTCACTGAGTGTGTATGTGCTTTTCTTCTATAGCCTCTA 990
DB 1062 CAGAAAACAAAGCTTTTCACTGAGTGTGTATGTGCTTTTCTTCTATAGCCTCTA 1121
QY 991 GATCTATGTCTCAGAAAGCATTTTAAGCAATTCCTTCATCATGAGGTCAATACAC 1050
DB 1122 GATCTATGTCTCAGAAAGCATTTTAAGCAATTCCTTCATCATGAGGTCAATACAC 1181
QY 1051 GAGTGTGCTTCTGCTGCTGTATGAGAAGAGCTCTGATATCTTCAAGTCTCAACAG 1110
DB 1182 GAGTGTGCTTCTGCTGCTGTATGAGAAGAGCTCTGATATCTTCAAGTCTCAACAG 1241
QY 1111 TCCCTTGCCCTCATCTATCAAAACATCTGCAATCCCGCTCATATTTGACCTT 1170
DB 1242 TCCCTTGCCCTCATCTATCAAAACATCTGCAATCCCGCTCATATTTGACCTT 1301
QY 1171 GTGGCTAATGAATCTTCAATGACAGCACTCCCTGACATGATCAAGTCTTCTG 1230
DB 1302 GTGGCTAATGAATCTTCAATGACAGCACTCCCTGACATGATCAAGTCTTCTG 1361
QY 1231 GACTTGTCTGGAAGATGTCTTTGTGTGTCTCTGACATGATCAAGTCTTCTG 1290
DB 1362 GACTTGTCTGGAAGATGTGTCTTTGTGTGTCTCTGACATGATCAAGTCTTCTG 1421
QY 1291 GATGCTGTGACACTGTCTTACTTATGATGTTTGGACCGGCTCAAGTCTTGAAGAC 1350
DB 1422 ----- 1421
QY 1351 ACGAAGCCGCTTTTGTCAAAAGCCGACACGCTGATGATGCTGTTGTGTGTGT 1410
DB 1422 ----- 1421
QY 1411 GCCTTCCGAAGGGGGACATTTGTTATGTGTGAAGAGACCCACGAGAGAGAAATTCTG 1470
DB 1422 -----GAAGAGACCAAGAGAGAGAAATTCTG 1451
QY 1471 AGCCGAAAGATGATGAATACTGGCTACTTGTCTGGAACCGGGAATCTTAATGGGAC 1530
DB 1452 AGCCGAAAGATGATGAATACTGGCTACTTGTCTGGAACCGGGAATCTTAATGGGAC 1511
QY 1531 GACCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1590
DB 1512 GACCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1571
QY 1591 ATGAGCTTCGAGAGAGACTCAAAAGAACCGGGGTGATTTTGTGACCAACACATCCCC 1650
DB 1572 ATGAGCTTCGAGAGAGACTCAAAAGAACCGGGGTGATTTTGTGACCAACACATCCCC 1631
QY 1651 CTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1710
DB 1632 CTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1691
QY 1711 CTCCTCAGGCTTTCTTTTCTTTTGTGTCTCTTGA 1746
DB 1692 CTCCTCAGGCTTTCTTTTCTTTTGTGTCTCTTGA 1727

```

RESULT 11
 ADF50146
 ID ADF50146 standard; cDNA; 2145 BP.
 XX
 AC ADF50146;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE Cat cauxin encoding cDNA SEQ ID NO:3.
 XX
 KW cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.
 XX
 OS Felis catus.
 XX
 PN JP2003250575-A.
 XX

PD 09-SEP-2003.
XX 04-MAR-2002; 2002JP-00057908.
XX 04-MAR-2002; 2002JP-00057908.
XX (TOHO-) TOHOKU TECHNOARCH KK.
XX MPI: 2004-002277/01.
DR P-PSDB; ADP50147.
XX
PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
and for diagnosing cat kidney disease.
XX
PS Claim 6; SEQ ID NO 3; 33pp; Japanese.
XX
CC The present sequence encodes a cat cauxin protein (I) or its salt, which
CC is cat kidney disease marker. Also described: (1) a partial peptide (II)
CC of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
CC comprising (III); (4) a transformed host (V) comprising (III) or (IV);
CC (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
CC couples specifically with (I) or (II); (7) diagnosing cat kidney disease
CC which involves measuring (I) quantitatively, and where reduction of
CC amount of (I) indicates presence of the disease; (8) a cat kidney disease
CC diagnostic agent comprising (I) labelling agent, a reagent which measures
CC the biological activity of urinary (I) or (VI); and (9) a cauxin
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.
XX
XX Sequence 2145 BP; 444 A; 581 C; 550 G; 564 T; 0 U; 6 Other;

Query Match 62.9%; Score 1099; DB 12; Length 2145;
Best Local Similarity 80.1%; Pred. No. 8.6e-314;
Matches 1329; Conservative 0; Mismatches 323; Indels 8; Gaps 3;

QY 91 GGGCTTGTCTGTAAGGCGCACAGAGAACACACAGGCTGGATGATTCAGGGCAAGCA 150
DB 238 GGGCGAGCTGCTGAGAGCAACAGTAGAGAGCACAGGCTGGATGATTCAGGGCAAGCA 297
QY 151 GTCACTGCTGGGAAGCGCTGCTGTAAGCGTTCCTCGAGTCCCTTCTGCT 210
DB 298 ACGACTGACTGGGAAGCACGCTGCTGTAAGCGTTCCTCGAGTCCCTTCTGCT 357
QY 211 CCCCCTGGGATCCCTGCGATTTACGAACCCGAGAGCTGATCGCCCTGGGATTAATTG 270
DB 358 CTTCTAGAGGCGCTGCGATTTAAGCAACCAAGCGCTGCTGCTCGGGGAATGACTTC 417
QY 271 CGAGAGCGACCTCTACCTTAATTTGCTGCTCGAAGCTCAGAGTGGCTCTTAAT 330
DB 418 CGAATGCGACATCTTACCTTAATTTGCTGCTCGAAGCTCAGAGTGGCTCTTAAT 477
QY 331 CAACAGATGCTCAAGTGCATTAACCGAATTCGAGAGTGAAGAGCTGCTCAACG 390
DB 478 CAACAGTTCGAAGTGCATTAACCGAATTTGAAGGTCGGAAGACTGCTGAACCT 537
QY 391 AACATCTATGCGCTGCCACGCGCGATACAGGCTCAAGCTCCCGCTTGGTGTGTTTC 450
DB 538 AACATCTATGCGCGAGCGCGCGATACAGGCTCAAGCTCCCGCTTGGTGTGTTTC 597
QY 451 CGAGAGGTCCTTCAAGACTGAGCTGAGCTCAATCTTATGAGTCCGCTGCTGCC 510
DB 598 CCGCGGGGCTGCTTCAAGATGAGGCTCAAGCTTCTCTTCAAGTGGCTGCTGCC 657
QY 511 TATAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
DB 658 TACGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 571 ACATGAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630

DB 718 ACAGGGATGAGATGCCCGGGGAACTGGGCTTGTGAGCAAGATGGCTCCCTAAC 777
QY 631 TGGGTCCAGGAAGATATGAGTCTTGGTGGGAGCCCAAGCTGTGATCATCTTTGGC 690
DB 778 TGGGTCCAGGAAGATATGAGTCTTGGTGGGAGCCCAAGCTGTGATCATCTTTGGC 837
QY 691 GAGTCCGGGAGCCATAGTGTCTTATCTTATCTGCTCCCATGGCCAAAGCTTTA 750
DB 838 GAGTCCGGGAGCCATAGTGTCTTATCTTATCTGCTCCCATGGCCAAAGCTTTA 897
QY 751 TTCCCAAGCCATATGAGAGTGGGTGGCCATCATCTTACCTGAG--GCCAT 807
DB 898 TTCCCAAGCCATATGAGAGTGGGTGGCCATCATCTTACCTGAGAGCCCT 957
QY 808 GATTATGAGAGAGTGAAGCTGAGAGTGTGCAATTTCTGATGAACAATGCTCA 867
DB 958 GGTATGAGAGAGAGAGATTTGAGAGTGTGAGCGATCTGATGTGCAATGCT 1017
QY 868 GACTCTGAGGCGCTGCTGAGTGTGCTGAGCAAAACCTCCAGAGAGCTGCTGACCTTC 927
DB 1018 GACTCTGAGGCGCTGCTGAGTGTGCTGAGCAAAACCTCCAGAGAGTGTGAGCAATC 1077
QY 928 AGCCAGAAACAAAGTCTTTCACTGAGTGTGATGAGTCTTCTTATAGCT 987
DB 1078 AGCAAGAACTCAAGTTTCACTGAGTGTGATGAGTCTTCTTATAGCT 1137
QY 988 CTAGATCTATGTCTCAAGAAAGATTTAAGCAATTCCTCATATGAGCTCAATAC 1047
DB 1138 GTAGGCTTATGATCAAAAGATTTAAGTCTTCTTATATGAGCTCAATAC 1197
QY 1048 CACGAGTGTGCTTCTGCTGCTGATGAGAGCTGCTGAGATCTGAGTGTCTCAAC 1107
DB 1198 CACGAGTGTGCTTCTGCTGCTGATGAGAGCTGCTGAGATCTGAGTGTCTCAAC 1254
QY 1108 AAGTCCCTTGGCTTCATCTGATGATCAAAACATCTGCAATCCGCTCAATTTGCA 1167
DB 1255 AAGTCTGAGGCGCTGCTGATGATCAACAGTTCCTGAATATTCACCAAGATTTGCA 1314
QY 1168 CTGTGCTTATGATTAATCTTCAAGCAAGCAATCCCTGATGAGTGTGCTGCT 1227
DB 1315 CTGTGCTTATGATTAATCTTCAAGCAAGCAATCCCTGATGAGTGTGCTGCT 1374
QY 1228 CTGAGCTTGTGAGATGTGTTCTTGTGCTGCTGATGATGATGATGATGATGATGAT 1287
DB 1375 CTGAGCTTGTGAGATGTGTTCTTGTGCTGCTGATGATGATGATGATGATGATGAT 1434
QY 1288 AGAGATGCTGTGCACTGCTTATCTATGAGTGTGCAACCGGCTTCACTGCTTGA 1347
DB 1435 AGAGATGCTGTGCACTGCTTATCTATGAGTGTGCAACCGGCTTCACTGCTTGA 1494
QY 1348 GACAGAGAGCGGCTTGTGCTCAAGCGACACAGCTGATGAGTGTGCTTGTGCT 1407
DB 1495 GACAGAGAGCGGCTTGTGCTCAAGCGACACAGCTGATGAGTGTGCTTGTGCT 1554
QY 1408 GGTGCTTCTGAGAGGGGAGCATTTGTATGTTGAGAGGACCGAGAGAGAGAGTGA 1467
DB 1555 GGTGCTTCTGAGAGGGGAGCATTTGTATGTTGAGAGGACCGAGAGAGAGAGTGA 1614
QY 1468 CTGAGCCGAGAGATGATGAAATTAATGAGCTTCTGAGACCGGAAATCTTAATGG 1527
DB 1615 CTGAGCCGAGAGATGATGAGTGTGCTGAGCTTCTGAGACCGGAAATCTTAATGG 1674
QY 1528 AACGACCTGTCTGTGCTGAGCTTATATCTGATGAGCAATCTCACTGAGCTTGG 1587
DB 1675 GAGGCTGTGCTGTGCTGAGCTTATATCTGATGAGCAATCTCACTGAGCTTGG 1734
QY 1588 AACGACCTGTGCTGAGCTTATATCTGATGAGCAATCTCACTGAGCTTGG 1646
DB 1735 AGTGTGAGCTGTGAGCAAGAACTGAGAGAGCAAGAGTGTGAGTGTGAGTGTGAGT 1794
QY 1647 -CCCCCTGATCTGTCTGCTGCTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1705
DB 1795 GTCCCTGATACCCCGCACTGAGAGGCTTCCCAAGTCTGCTTNTCCCTTACCTGCC 1854

QY 1706 TCTCTCTCCAGCCCTTTCTTTCTTTGTGCTCTTG 1745
DB 1855 TTCTTTGTCTCCGCTGGCTTTTCTTTCTCTCATG 1894

RESULT 12
ID ADF50144 standard; cDNA; 1629 BP.
AC ADF50144;
XX

DT 12-FEB-2004 (first entry)

DE Cat cauxin encoding cDNA SEQ ID NO:1.

KM cat; cauxin; cat kidney disease marker; kidney disease; gene; ss.

OS Feline catub.

PN JP2003250575-A.

XX 09-SEP-2003.

PF 04-MAR-2002; 2002JP-00057908.

PR 04-MAR-2002; 2002JP-00057908.

PA (TOHO-) TOHOKU TECHNORACH KK.

DR WPI; 2004-002277/01.

XX P-PSDB; ADF50145.

PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,
DB and for diagnosing cat kidney disease.

XX Claim 6; SEQ ID NO 1; 33pp; Japanese.

XX The present sequence encodes a cat cauxin protein (I) or its salt, which
CC is cat kidney disease marker. Also described: (1) a partial peptide (II)
CC of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector (IV)
CC comprising (III); (4) a transformed host (V) comprising (III) or (IV);
CC (5) producing (I) or (II) by culturing (V); (6) an antibody (VI) which
CC couples specifically with (I) or (II); (7) diagnosing cat kidney disease
CC which involves measuring (I) quantitatively, and where reduction of
CC amount of (I) indicates presence of the disease; (8) a cat kidney disease
CC diagnostic agent comprising (I) labelling agent; a reagent which measures
CC the biological activity of urinary (I) or (VI); and (9) a cauxin
CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly. (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
XX blood testing.

XX Sequence 1629 BP; 345 A; 438 C; 441 G; 405 T; 0 U; 0 Other;

Query Match 60.9%; Score 1063.2; DB 12; Length 1629;
Best Local Similarity 80.9%; Pred. No. 3e-303;
Matches 1265; Conservative 0; Mismatches 293; Indels 6; Gaps 2;

QY 91 GGGGCTTCTGCTGAAGGCGCAGAGAAACAGGCTGGATGATTCAGGGCAACAA 150
DB 64 GGGGCACTGCTGATGACCAAGAGAGACACAGGCTGGATGATTCAGGGCAACAA 123
QY 151 GTGCTGCTGCTGAAGGCGCAGAGAAACAGGCTGGATGATTCAGGGCAACAA 127
DB 124 ACACCTTACTGAGAGACACGATGCTGTAACATGTTCTCTGGATCCCTTAAGTGA 183
QY 211 CCCCCGCTGGATCCCTGCGATTTACGAACCCGAGCCTGCATGCGCTGGATTAATCTG 270
DB 184 CTTCCTTAAGGGCCCTGCGATTTAAGCAACCAAGCTCTCTGCGGGGAATGACTTC 243

QY 271 CGAAGGCCACCTCCTAACCTTAATTTGTGCTCCAGAACTCAGAGTGTGCTCTTAGAT 330
DB 244 CGAAATGCCACATCTTACCTTAATTTATGCTTCCAGGACTTAAGTGTGCTGTCTTAT 303
QY 331 CAACACATGCTCAAGTGTGATTTACCCGAATTTGAGTGTCAAGAACTGCTCTTACTG 390
DB 304 CAACAGTGTCTCAAGTGTGATTTACCCGAATTTGAGTGTCAAGAACTGCTGTACTT 363
QY 391 AACATCTATGAGCGCTGCCCAAGCGGATACAGAGCTCAAGCTCCCGTGTGAGTGTG 450
DB 364 AACATCTATGAGCGCGAGGCCCAAGCGGATACAGAGCTCAAGCTCCCGTGTGAGTGTG 423
QY 451 CCAGAGGTGCTTCAAGACTGAGCTTCAAGCTTCAAGCTTCAAGTGTGAGTGTGAGTGTG 510
DB 424 CCGGGGCTGCTTCAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 483
QY 511 TATGAGAGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 570
DB 484 TACGAGAGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 543
QY 571 ACATGAGATCAGATGCTCCGAGGAACTGAGCTTCAAGGACAGTGTGCTGTGCTGTC 630
DB 544 ACAGGGATGAGATGCGCGGGGAACTGAGCTTGTGAGACAGAGTGTGCTGTGCTGTC 603
QY 631 TGGGTCCAGAAAGAACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 690
DB 604 TGGGTCCGAGCAACATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 663
QY 691 GAGTCCGAGGAGCATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 750
DB 664 GAGTCCGAGGAGCATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 723
QY 751 TTCCAGAAAGCATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 807
DB 724 TTCCAGAAAGCATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 783
QY 808 GATATGAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 867
DB 784 GATATGAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 843
QY 868 GACTGTGAGGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 927
DB 844 GACTGTGAGGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 903
QY 928 AGCCAGAAAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 987
DB 904 AGCCAGAAAGAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 963
QY 988 CTGATCTATTTCTCAGAAAGATTTAAGCAATCTTCCATCTGAGTGTGAGTGTGAGT 1047
DB 964 GTAGCCCTATTTCTCAGAAAGATTTAAGCAATCTTCCATCTGAGTGTGAGTGTGAGT 1023
QY 1048 CACGAGTGTGCTTCTGCTGCTGCTATGAGAGAGCTCTGAGATCTCTGAGTGTGAGT 1107
DB 1024 CACGAGTGTGCTTCTGCTGCTGCTATGAGAGAGCTCTGAGATCTCTGAGTGTGAGT 1080
QY 1108 AGTCCCTTCCCTCCTCATCTGATGATCAAAACATCTGCAATCTCCGCTGATTTGCA 1167
DB 1081 AGTCCCTTCCCTCCTCATCTGATGATCAAAACATCTGCAATCTCCGCTGATTTGCA 1140
QY 1168 CTGTGCTTATGAAATCTTCAAGCAAGCACTCCCTCAAGTGTGAGTGTGAGTGTGAGT 1227
DB 1141 CTGTGCTTATGAAATCTTCAAGCAAGCACTCCCTCAAGTGTGAGTGTGAGTGTGAGT 1200
QY 1228 CTGAGCTTGTGAGATGTGTTCTTTGTGCTCCCTGCACTGATCAAGCTGATATCAC 1287
DB 1201 CTGAGCTTGTGAGATGTGCTTTGTGCTCCCTGCGGTGATCAAGCTGATATCAT 1260
QY 1288 AGAGATGTGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1347
DB 1261 AGAGATGTGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1348 GACAGAGAGCCGCTTTGTCTAAAGCCGACGCTGATGAGTGTGATGAGTGTGATGAGT 1407

Db 1321 GACACGAGGCCGACTTTCGTGAAACCCGATCACTGATGAATCCGCTTCTTGGGA 1380
 Qy 1408 GGTGCTTCTCGAAGGCGGACATTGTATGTTGCGAAGAGCCACGAGAGAGAGAGTTA 1467
 Db 1381 GGTGCTTCTCGAAGGCGGACATTGTATGTTGCGAAGAGCCACGAGAGAGAGAGTTG 1440
 Qy 1468 CTGAGCCGGAAGATGATGAAATGATGAGGCTACCTTGTGTCGACCGGGAATCTTAAGG 1527
 Db 1441 CTGAGCAGAGAGATGATGAGGCTACCTTGTGTCGACCGGGAATCTTAAGG 1500
 Qy 1528 AACGACCTGTCTGTGCGCAGCTTATATCTGACTGAGCAGTACCTCCAGCTGAGCTTG 1587
 Db 1501 GAAGGTGTCCTCTGTGCGCAGCTTATATCTGACTGAGCAGTACCTCCAGCTGAGCTTG 1560
 Qy 1588 AACATGACCTGTGCGCAGCAGATCTCAAAGAACCCGGGTGAGATTTTGGACCGACCATTC 1647
 Db 1561 AGTGTGACCTGTGCGCAGCAGATCTCAAAGAACCCGGGTGAGATTTTGGATGATACCAT 1620
 Qy 1648 CCCC 1651
 Db 1621 GTCC 1624

RESULT 13.
 ID ABQ86171 standard: DNA; 1071 BP.
 AC ABQ86171;
 DT 10-SEP-2002 (first entry)
 DE Novel human gene. SEQ ID 42.

Human; cytostatic; vulnerary; antiarteriosclerotic; antiparkinsonian;
 neurotropic; neuroprotective; immunosuppressive; haemostatic;
 antiinflammatory; cardiant; antitumor; virucide; antihypertoid;
 cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
 wound healing disorders; atherosclerosis; Parkinson's disease;
 Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
 inflammation; neoplastic disease; nervous system disorder;
 cardiovascular disorders; pancreatitis; respiratory disorder;
 hyperproliferation; systemic autoimmune disease; hyper-immunity;
 developmental abnormality; gastrointestinal ulceration; neuropathy;
 haematological disease; metabolic disease; sperm dysfunction;
 thyroid disorder; hypothyroidism; brain damage; colitis;
 cone photo- transduction deficiency; neurological disease; stroke;
 angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
 trachea; thymus; lymph node; muscular system; obesity; anorexia;
 growth abnormality; precocious puberty; gene; ss.

OS Homo sapiens.
 XX WO200250105-A1.
 XX 27-JUN-2002.
 PD 17-DEC-2001; 2001WO-US049232.
 XX 19-DEC-2000; 2000US-0256710P.
 PR 20-DEC-2000; 2000US-0257048P.
 PR 09-JAN-2001; 2001US-0260483P.
 PR 30-JAN-2001; 2001US-0268922P.
 PR 06-FEB-2001; 2001US-0266797P.
 PR 19-MAR-2001; 2001US-0276988P.
 PR 04-APR-2001; 2001US-0281535P.
 PR 08-MAY-2001; 2001US-0289622P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PA (GLAX) GLAXO GROUP LTD.
 XX Aggarwal P, Birkeland M, Cogswell JP, Kahnlick KF, Lai Y;

PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
 XX WPI: 2002-508784/54.
 DR P-PsDB; ABP61006.
 XX
 PT Secreted proteins and polynucleotides useful as vaccines for preventing
 PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
 PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
 XX
 PS Claim 2(a); Page 252; 335pp; English.

The invention relates to an isolated polypeptide with signal sequences
 CC which allow it to be secreted extracellularly or membrane associated. The
 CC activity of polypeptides of the invention may be described as,
 CC cytostatic, vulnerary, antiarteriosclerotic, antiparkinsonian, neurotropic,
 CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
 CC cardiant, antitumor, virucide, antihypertoid, cerebroprotective, anorectic,
 CC and metabolic. Polypeptides and polynucleotides of the invention are
 CC useful in the treatment, or as a vaccine in the prevention of, cancer,
 CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
 CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
 CC inflammation, neoplastic diseases, nervous system related disorders and
 CC cardiovascular disorders, pancreatitis, respiratory disorder.
 CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
 CC developmental abnormality, gastrointestinal ulceration, neuropathy,
 CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
 CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
 CC transduction deficiency, neurological diseases, stroke, angiogenesis,
 CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities, and alleviation of precocious puberty. The
 CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's
 CC of the invention

SQ Sequence 1071 BP; 242 A; 286 C; 254 G; 289 T; 0 U; 0 Other;

Query Match 58.5%; Score 1021.4; DB 6; Length 1071;
 Best Local Similarity 99.9%; Pred. No. 5,5e-291;
 Matches 1022; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 724 ATACTGTCTCCCATGCGCAAAAGGCTTATTCACAAAGCATCATGAGAGTGGGTGGCC 783
 Db 49 ATACTGTCTCCCATGCGCAAAAGGCTTATTCACAAAGCATCATGAGAGTGGGTGGCC 108
 Qy 784 ATCATCCCTTACCTGAGAGCCCAAGATATGAGAGAGTGAAGTCTCAAGTGGTTGCA 843
 Db 109 ATCATCCCTTACCTGAGAGCCCAAGATATGAGAGAGTGAAGTCTCAAGTGGTTGCA 168
 Qy 844 CATTCTGTGATTAACATGCGTCAAGACTCTGAGGCGCTGAGGTGCTGAGACAATA 903
 Db 169 CATTCTGTGATTAACATGCGTCAAGACTCTGAGGCGCTGAGGTGCTGAGACAATA 228
 Qy 904 CCTTCGAAGAGCTGCTGAGCCCTCAGCCAGAAAACAAAGTCTTTGACTGAGTGGTTGAT 963
 Db 229 CCTTCGAAGAGCTGCTGAGCCCTCAGCCAGAAAACAAAGTCTTTGACTGAGTGGTTGAT 288
 Qy 964 GGTGCTTCTTCTTCAATGAGGCTTATGATCTATGTTCTCAGAAAGCATTTAAAGCAATT 1023
 Db 289 GGTGCTTCTTCTTCAATGAGGCTTATGATCTATGTTCTCAGAAAGCATTTAAAGCAATT 348
 Qy 1024 CCTTCATCATGCGAGGATTAACCAAGAGTGTGCTTCGAGTCTCATGAAAGAGGCT 1083
 Db 349 CCTTCATCATGCGAGGATTAACCAAGAGTGTGCTTCGAGTCTCATGAAAGAGGCT 408
 Qy 1084 CCTGAGATCTCAGTGGCTCCCAAGAGTCCCTTGCCCTCACTGTATCAAAAACATCTCTG 1143
 Db 409 CCTGAGATCTCAGTGGCTCCCAAGAGTCCCTTGCCCTCACTGTATCAAAAACATCTCTG 468
 Qy 1144 CACATCCCGCTCAGATATTTGCACTTGTGGCTAATGAATCTTCCATGAGCAAGACATCC 1203
 Db 469 CACATCCCGCTCAGATATTTGCACTTGTGGCTAATGAATCTTCCATGAGCAAGACATCC 528
 Qy 1204 CTGACTGAATCCGAGACAGTCTTCTGAGCTTGAGAGATGTGTTCTTTGTGTCTCCT 1263

```
Db 529 CTGACGAAATCCGAGCAGCTTCTGGACTTGCTGGAAATGTTCTTTGGGCTCT 588
Qy 1264 GCACTGATCAGACTGATATCAAGAGATGCTGGACCTGTCTAATGAGATT 1323
Db 589 GCACTGATCAGACTGATATCAAGAGATGCTGGACCTGTCTAATGAGATT 648
Qy 1324 CGGACACCGGCTCAGTGTCTTGAAGACAAGACCGGCTTTTCTCAAGCCGACCGCT 1383
Db 649 CGGACACCGGCTCAGTGTCTTGAAGACAAGACCGGCTTTTCTCAAGCCGACCGCT 708
Qy 1384 GATGAAGTCGGCTTGTGTGCTGGTGTGCTTCTGAAGGGGACATGTATGTTGAA 1443
Db 709 GATGAAGTCGGCTTGTGTGCTGGTGTGCTTCTGAAGGGGACATGTATGTTGAA 768
Qy 1444 GGAGCCAGGAGGAGGAGAAATTAAGAGCCGGAAGATGATGAATATCGGCTACTTT 1503
Db 769 GGAGCCAGGAGGAGGAGAAATTAAGAGCCGGAAGATGATGAATATCGGCTACTTT 828
Qy 1504 GCTCGAACCAGGAAATCTAATGGAACGACCTGTCTGTGGCCAGCTTATATCTGACT 1563
Db 829 GCTCGAACCAGGAAATCTAATGGAACGACCTGTCTGTGGCCAGCTTATATCTGACT 888
Qy 1564 GAGCAGTACTCTCAGCTGATCTTGAACATGAGCCTTGGACAGAGACTCAAGAACCGCG 1623
Db 889 GAGCAGTACTCTCAGCTGATCTTGAACATGAGCCTTGGACAGAGACTCAAGAACCGCG 948
Qy 1624 GTGATTTTGGACAGACATCCCTGATCTGTGCTGCTCGACATGCTCAAGT 1683
Db 949 GTGATTTTGGACAGACATCCCTGATCTGTGCTGCTCGACATGCTCAAGT 1008
Qy 1684 CCTTTTCTCTTAACTTCTCTCTCTCTCTCAAGCCTTTCTTTTCTTTGCTCTCT 1743
Db 1009 CCTTTTCTCTTAACTTCTCTCTCTCTCTCAAGCCTTTCTTTTCTTTGCTCTCT 1068
Qy 1744 TGA 1746
Db 1069 TGA 1071

RESULT 14
ADCS5523
ID ADCS5523 standard; cDNA, 1244 BP.
XX
AC ADCS5523;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human carboxylase 24.64 encoding sequence.
XX
KW human carboxylase-24.64; primary hypertension; digestive ulcer;
KM nephrotic; bronchial asthma; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 236..910
FT FT /*Cag= a
XX
PN CN1382799-A.
XX
PD 04-DEC-2002.
XX
PF 26-APR-2001; 2001CN-00112736.
XX
PR 26-APR-2001; 2001CN-00112736.
XX
PA (BIOW-) BIOWINDOM GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2003-269506/27.
XX
DR P-PSDB; ADCS5524.
DR
```

```
XX Polypeptide-human carboxylase-24.64 and polynucleotide for coding it.
PT
XX
XX Claim 6; SEQ ID NO 1; 31p; Chinese.
XX
XX The present invention relates to a polypeptide-human carboxylase-24.64,
CC the polynucleotide for coding it, the process for preparing the
CC polypeptide by DNA recombination, the application of the polypeptide in
CC treating diseases such as primary hypertension, digestive ulcer,
CC nephrotic, bronchial asthma, tremor, etc, the antagonist of the polypeptide
CC and its medical action, and the application of the polynucleotide are
CC new. The present sequence represents human carboxylase 24.64 encoding
CC sequence.
XX
SQ Sequence 1244 BP; 301 A; 314 C; 271 G; 358 T; 0 U; 0 Other;
Query Match 51.8%; Score 905.2; DB 10; Length 1244;
Best Local Similarity 99.7%; Pred. No. 1.4e-256;
Matches 907; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 837 GGTTCGACATTTCTGTGTAAACAATGCGTCAGACTTGAAGCCCTGTGAGTCTGAG 896
Db 1 GGTTCGACATTTCTGTGTAAACAATGCGTCAGACTTGAAGCCCTGTGAGTCTGAG 60
Qy 897 GACAAAACCTCCAAAGAGCTGAGCCCTCAGCCAGAAAACAAGTCTTCACTCGAGT 956
Db 61 GACAAAACCTCCAAAGAGCTGAGCCCTCAGCCAGAAAACAAGTCTTCACTCGAGT 120
Qy 957 GGTTCGATGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1016
Db 121 GGTTCGATGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 180
Qy 1017 AGCAATTCCTTCCATCTGAGATCAATACAGAGATGCTTCTTCTTCTTCTTCTTCT 1076
Db 181 AGCAATTCCTTCCATCTGAGATCAATACAGAGATGCTTCTTCTTCTTCTTCTTCT 240
Qy 1077 GAGAGCTCTGAGATCTCTCAGTGGCTCAACAAGTCCCTGACCTCATGATGAA 1136
Db 241 GAGAGCTCTGAGATCTCTCAGTGGCTCAACAAGTCCCTGACCTCATGATGAA 300
Qy 1137 CATCTGCAATCCCGCTCAGTATTTGACCTTGTGCTTAAATGAAATCTTCATGACAA 1196
Db 301 CATCTGCAATCCCGCTCAGTATTTGACCTTGTGCTTAAATGAAATCTTCATGACAA 360
Qy 1197 GCACTCCCTGACGAATATCCGAGACAGTCTTCTGAGCTTCTGAGATGTTCTTTGT 1256
Db 361 GCACTCCCTGACGAATATCCGAGACAGTCTTCTGAGCTTCTGAGATGTTCTTTGT 420
Qy 1257 GGTCCCTGCACTGATCAAGCTGATATCAAGAGATGCTGAGTCACTGTCTACTTCT 1316
Db 421 GGTCCCTGCACTGATCAAGCTGATATCAAGAGATGCTGAGTCACTGTCTACTTCT 480
Qy 1317 TGAATTTTGGACACCGGCTCAGTGTCTTGAAGACAAGACCGGCTTTTGTCAAGCCGA 1376
Db 481 TGAATTTTGGACACCGGCTCAGTGTCTTGAAGACAAGACCGGCTTTTGTCAAGCCGA 540
Qy 1377 CCAAGCTGATGAAGTCCGCTTGTGTTGCTGAGTGGCTTCTGAAGGGGGAATGTAT 1436
Db 541 CCAAGCTGATGAAGTCCGCTTGTGTTGCTGAGTGGCTTCTGAAGGGGGAATGTAT 600
Qy 1437 GTTCGAAGAGCCAGGAGGAGAGAAATTAAGTCAAGCCGGAAGATGAAATATCTGGCC 1496
Db 601 GTTCGAAGAGCCAGGAGGAGAGAAATTAAGTCAAGCCGGAAGATGAAATATCTGGCC 660
Qy 1497 TACCTTGTCTGAACCGGGAATCTTAATGGAACGACCTGTCTGTGGCCAGCTTATTA 1556
Db 661 TACCTTGTCTGAACCGGGAATCTTAATGGAACGACCTGTCTGTGGCCAGCTTATTA 720
Qy 1557 TCTGACTGAGAGTACTCTCAGCTGAGCTTGAACATGAGCCTCGGACAGAGACTCAAGA 1616
Db 721 TCTGACTGAGAGTACTCTCAGCTGAGCTTGAACATGAGCCTCGGACAGAGACTCAAGA 780
Qy 1617 ACCGGGGGTGATTTTGTGACCAAGACCATCCCTGATCTGTGCTCGACATGCT 1676
```

```

DB 761 ACCGGGGGGGATTTTGGACGACGACATCCCGCGATCCCGTGCCTCGACATGCT 840
QY 1677 CCACAGTCTCTTTCTTCTTAACTTCTCTCTCTCTCCACGCTTTCTTTTCTTTG 1736
DB 841 CCACAGTCTCTTTCTTCTTAACTTCTCTCTCTCTCCACGCTTTCTTTTCTTTG 900
QY 1737 TGCTCCTTGA 1746
DB 901 TGCTCCTTGA 910

RESULT 15
AAD33344
ID AAD33344 standard, cDNA, 921 BP.
XX
AC AAD33344;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human carboxylesterase-like enzyme cDNA.
XX
KM Human; carboxylesterase-like enzyme; organophosphorus intoxication; gene;
XX osteopathic; gene therapy; osteoporosis; antisense therapy; cytostatic;
XX detoxifying agent; Paget's disease; bone implant degradation; cancer;
XX dental implant; enzyme; gene expression; ss.
XX Homo sapiens.
XX
PH Key Location/Qualifiers
FT 1..921
CDS /*tag= a
FT /product= "Human carboxylesterase-like enzyme"
FT
PN WO200206454-A2.
XX
PD 24-JAN-2002.
XX
PR 10-JUL-2001; 2001WO-EP007919.
XX
PR 17-JUL-2000; 2000US-0218564P.
XX (PARB ) BAYER AG.
XX
PI Xiao Y;
XX
DR WPI; 2002-195808/25.
DR P-PSDB; AAE20909.
XX
PT Novel human carboxylesterase-like enzyme polypeptide, regulators of which
PT are useful for preventing and treating organophosphorus intoxication,
PT cancer and osteoporosis.
XX
PS Claim 1; Fig 4; 92pp; English.
XX
XX The invention relates to a purified human carboxylesterase-like enzyme
XX polypeptide. Carboxylesterase-like enzyme and its DNA are useful for
XX screening for agents which decrease or modulate the activity of
XX carboxylesterase-like enzyme polypeptide. Carboxylesterase-like enzyme is
XX useful for treating a carboxylesterase-like enzyme dysfunction related
XX disease, such as organophosphorus intoxication, cancer and osteoporosis.
XX Compounds that increase the ability of human carboxylesterase-like enzyme
XX to bind to organophosphorus compounds are useful as detoxifying agents.
XX Carboxylesterase-like enzyme agonists and antagonists are useful for
XX treating osteoporosis. Paget's disease and degradation of bone implants,
XX particularly dental implants. Carboxylesterase-like enzyme is useful in
XX diagnostic assays for detecting diseases and abnormalities or
XX susceptibility to diseases or abnormalities related to the presence of
XX mutations in the nucleic acid sequences which encode the enzyme. The
XX coding sequence of carboxylesterase-like enzyme polynucleotide is useful
XX in gene therapy and for generating antisense oligonucleotides or
XX ribozymes which specifically bind to mRNA transcribed from
XX carboxylesterase-like enzyme DNA. These antisense oligonucleotides are

```

```

CC useful for modulating carboxylesterase-like enzyme gene expression. The
CC present sequence is human carboxylesterase-like enzyme cDNA
XX
SQ Sequence 921 BP; 206 A; 238 C; 237 G; 240 T; 0 U; 0 Other;
Query Match 36.3%; Score 634.6; DB 6; Length 921;
Best Local Similarity 86.3%; Pred. No. 1.4e-176;
Matches 743; Conservative 0; Mismatches 49; Indels 69; Gaps 1;
QY 434 CCGTCTGTGTGTGTTCCTCCAGAGAGTGCCTTCAAGACGTGCTCAGCTCCATCTTGAG 493
DB 56 CAGTTTGTGTGTGTTCCTCCAGAGAGTGCCTTCAAGACGTGCTCAGCTCCATCTTGAG 115
QY 494 GGTCCGCCCTGCTGCTGCTATGAGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 553
DB 116 GGTCTCTCTGCTGCTGCTATGAGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 175
QY 554 TATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 613
DB 176 TATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 235
QY 614 AGGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 673
DB 236 AGGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 295
QY 674 CTGTGACCATCTTTGGCGAGTCCGGGGAGCCATAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 733
DB 296 CTGTGACCATCTTTGGCGAGTCCGGGGAGCCATAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 355
QY 734 CCATGGCCCAAGAGCTTATTCACAAAGCCATGAGAGTGGGTGGCCATCATCCTT 793
DB 356 CCATGGCCCAAGAGCTTATTCACAAAGCCATGAGAGTGGGTGGCCATCATCCTT 415
QY 794 ACCTGAAGCCCATGATTTATGAGAGAGTGAAGACCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 853
DB 416 ACCTGAAGCCCATGATTTATGAGAGAGTGAAGACCTGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 475
QY 854 GTTACAAAGCCGTCAAGCTCTGAGGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAAAG 913
DB 476 GGTACAAAGCCGTCAAGCTCTGAGGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAAAG 535
QY 914 AGCTGCTGACCTCTGAGCCAGAAACAAAGCTCTTCACTGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 973
DB 536 AGCTGCTGACCTCTGAGCCAGAAACAAAGCTCTTCACTGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 595
QY 974 TTCTTAATGAGCCCTGATCTTATGTCTCAAGAGCATTTTAAAGCATTTCTTCCATCA 1033
DB 596 TTCTTAATGAGCCCTGATCTTATGTCTCAAGAGCATTTTAAAGCATTTCTTCCATCA 655
QY 1034 TGGAGTCAATTAACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1093
DB 656 TGGAGTCAATTAACACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 694
QY 1094 TCAGTGTCTCAACAAAGTCCCTTGCCTTCATCTGATACAAACATCTGCACATCCGC 1153
DB 695 -----TGCATCTCTG 706
QY 1154 CTCAGTATTGCACTTGTGCTTAATGAACTTCCATGACAGACCTCCGACTGAA 1213
DB 707 CCAATTATTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 766
QY 1214 TCCGAGACAGCTTCTGAGCTTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1273
DB 767 TCCGAGACAGCTTCTGAGCTTGTGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 826
QY 1274 CAGCTCGATATCAAGAGATG 1294
DB 827 CAGCTCGATATCAAGAGATG 847

```

Search completed: June 14, 2005, 18:10:11
 Job time : 944.402 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 14, 2005, 23:03:41 ; Search time 127.014 Seconds
(without alignments)
10633.187 Million cell updates/sec

Title: US-10-023-515-3
Perfect score: 3170
Sequence: 1 atgcacacagsgactacttc.....ttctcttctgtcctctga 1746

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xln
-Q=/cgml1/USPTO.spool/6664091.runcat_14062005_133458_9325/app.query.fasta.1.4238
-DB=A_Geneseq_16Dec04 -OPMT=faetan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=2000000000
-USER=6664091.QCGN_1_1_299.gnumat_14062005_133458_9325 -NCPUR=6 -ICPU=3
-NO MMAP -LARGEQUERRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04:*

1: geneeegp1980s:*\n2: geneeegp1980s:*\n3: geneeegp2000s:*\n4: geneeegp2000s:*\n5: geneeegp2002s:*\n6: geneeegp2003as:*\n7: geneeegp2003bs:*\n8: geneeegp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3079	97.1	581	5	ABB79537 Human car
2	3079	97.1	581	8	ADQ89094 Human uro
3	2915	92.0	575	5	ABP61004 Human dia
4	2900	91.5	642	5	AAE25025 Human dr
5	2900	91.5	642	8	ABM84114 Human dia
6	2607.5	82.3	581	5	ABP61005 Human hum
7	2606	82.2	525	6	ADA54807 Human pro
8	2597.5	81.9	581	6	ABU54639 Human NOV
9	2588.5	81.7	618	5	ADR19663 Human dru
10	2454	77.4	469	7	ADB64065 Human pro

11	1985	62.6	542	8	ADP50145	Adf50145 Cat cauxi
12	1985	62.6	542	8	ADP50147	Adf50147 Cat cauxi
13	1780	56.2	356	5	ABP61006	Abp61006 Novel hum
14	1262.5	39.8	561	5	AAE20911	Aae20911 Rat carbo
15	1230	38.8	549	3	AAE58981	Aae58981 Breast an
16	1230	38.8	550	8	ADG98219	Adg98219 Human int
17	1230	38.8	559	4	AAAB31700	Aab31700 Protein e
18	1230	38.8	559	4	ABU53223	Abu53223 Human met
19	1230	38.8	559	6	ABR82851	Abbr82851 CES2 rela
20	1230	38.8	559	6	AAE33960	Aae33960 Human car
21	1230	38.8	559	7	ADD46521	Add46521 Human pro
22	1221	38.5	306	5	AAE20909	Aae20909 Human car
23	1213.5	38.3	561	7	ADD46519	Add46519 Rat Prote
24	1191	37.6	607	4	ABU53222	Abu53222 Human met
25	1186.5	37.4	554	8	AD179890	Ad179890 Mouse liv
26	1181	37.3	583	4	ABG10273	Abg10273 Novel hum
27	1179.5	37.2	571	3	AAV71107	Aav71107 Human Hyd
28	1179.5	37.2	571	4	AAU12442	Aau12442 Human PRO
29	1179.5	37.2	571	4	AAE04101	Aae04101 Human gen
30	1179.5	37.2	571	5	AAU83696	Aau83696 Human PRO
31	1179.5	37.2	571	5	ABR84949	Abbr84949 Human PRO
32	1179.5	37.2	571	5	ABG64341	Abg64341 Human alb
33	1179.5	37.2	571	6	ABO17886	Abol17886 Novel hum
34	1179.5	37.2	571	6	ABU69108	Abu69108 Human PRO
35	1179.5	37.2	571	6	ABU80843	Abu80843 Human PRO
36	1179.5	37.2	571	6	ABO33809	Abol33809 Novel hum
37	1179.5	37.2	571	6	ABU81140	Abu81140 Human PRO
38	1179.5	37.2	571	6	ABO19424	Abol19424 Human sec
39	1179.5	37.2	571	6	ABU66840	Abu66840 Human PRO
40	1179.5	37.2	571	6	ABU59921	Abu59921 Novel sec
41	1179.5	37.2	571	6	ABU69085	Abu69085 Human PRO
42	1179.5	37.2	571	6	ABO25111	Abol25111 Human sec
43	1179.5	37.2	571	6	ABU82152	Abu82152 Novel hum
44	1179.5	37.2	571	6	ABU67116	Abu67116 Human sec
45	1179.5	37.2	571	6	ABU81549	Abu81549 Human sec

ALIGNMENTS

RESULT 1	ABB79537 standard; protein; 581 AA.
ABB79537	
AC	ABB79537;
XX	
XX	23-SEP-2002 (first entry)
XX	
DB	Human carboxylesterase family member 53010.
XX	
KW	Carboxylesterase; enzyme; human; analgesic; nootropic; antiinflammatory;
KW	diagnosis; therapy.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Domain
FT	Region
FT	Active-site
FT	
FT	Location/Qualifiers
FT	1..26
FT	/label= Signal_peptide
FT	27..581
FT	/label= Mature_protein
FT	44..545
FT	/note= "carboxylesterase domain"
FT	125..135
FT	/note= "predicted carboxylesterase type-B signature"
FT	219..234
FT	/note= "predicted carboxylesterase type-B serine active site"
XX	
XX	WO200250256-A2.
XX	
PD	27-JUN-2002.
XX	
PF	18-DEC-2001; 2001WO-US049075.
XX	

PR 18-DEC-2000; 2000US-0256369P.
 PR 28-MAR-2001; 2001US-0279508P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ, Silos-Santiago I;
 XX
 DR WPI; 2002-547936/58.
 DR N-PSDB; ABRN4302.
 XX
 PT 53010 nucleic acids, useful for diagnosing and treating e.g. vascular
 PT diseases, autoimmune diseases, or neurodegenerative diseases, as
 PT surrogate markers, in tissue typing and chromosome mapping.
 XX
 PS Claim 5; Page 110-111; 122pp; English.

The present sequence is the protein sequence of human 53010, a novel protein containing the structural characteristics of a carboxylesterase family member. The invention provides 53010 nucleic acids, antisense molecules, expression vectors, host cells, transgenic animals, 53010 proteins, fusion proteins, antigenic peptides, anti-53010 antibodies and methods for detecting the presence of 53010 polypeptides or nucleic acids, of identifying a compound that binds to the 53010 polypeptide, and of modulating the activity of the polypeptide. The 53010 polypeptides and nucleic acids can act as novel diagnostic and therapeutic agents for controlling disorders involving aberrant or deficient hydrolysis of carboxylic esters. As 53010 mRNA is highly expressed in the central and peripheral nervous system, and its expression is regulated in some rodent pain models, 53010 molecules can also act as novel diagnostic targets and therapeutic agents for controlling neurological disorders, such as pain-related disorders. A claimed method of treating or preventing a disorder (especially a pain-related disorder) characterised by aberrant activity of a 53010-expressing cell involves administering a compound that modulates 53010 activity or expression

Sequence 581 AA:

Alignment Scores:
 Pred. No.: 3,21e-287 Length: 581
 Score: 3079.00 Matches: 581
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 97.13% Gaps: 0
 DB: 5 Indels: 0

US-10-023-515-3 (1-1746) x ABB79537 (1-581)

QY 1 ATGCACAGGAGCTTACTTCATCTGCTTACAGATGGTCTTTTCGATTCCTCCAGCC 60
 DB 1 MetProteinGlyLeuThrSerSerAlaSerGlnTrpCySPhePheLeuIleLeuGlnPro 20
 QY 61 CTGTTGGACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGGCCACAGAGAAC 120
 DB 21 LeuLeuGlnIleHisArgGlnTrpGlyLysThrGlyProSerAlaGlnGlyProGlnIleHisArg 40
 QY 121 ACCAGGCTGGGATGGATTCACAGGACAGACAGTCTGTGCTGGAGACCTGTGCTGTG 180
 DB 41 ThrArgLeuGlnGlyTrpIleGlnGlyLysGlnValThrValLeuGlnGlySerProValProVal 60
 QY 181 AACGTGTCTCTGGAGTCCCTTTGCTGCTCCCGCTGGGATCCCGGATTTATGGAAC 240
 DB 61 AsnValPheLeuGlnGlyValProPheAlaIleProProLeuGlnGlySerLeuArgPheThrAsn 80
 QY 241 CCGGACCTTCATGCGCTTGGGATTAATTGCGAGAACCACTTACCTTAATTTGTGC 300
 DB 81 ProGlnProAlaSerProTrpAspAsnLeuArgGlnIleThrSerTrpProAlaLeuGly 100
 QY 301 CTCGAGAACTCAGAGGCGCTGCTTATGATCAACACAGTCTCAAGTGCTTACCGGAAA 360
 DB 101 LeuGlnHisSerGlnValTrpLeuLeuAspGlnHisMetLeuValHisTrpTrpGly 120
 QY 361 TTGGAAGTGTCAAGAGCTGCTCTTACCTGAACATGTATGCGCTGCGCGGATTAACA 420

DB 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaSerThr 140
 QY 421 GACTCCAAAGTCCCGGTCTGTGTGTGTTCACAGAGTGCCTTCAAGACTGCTCAACC 480
 DB 141 GlySerIleValProValLeuValTrpPheProGlyGlyAlaPheLeuThrGlySerAla 160
 QY 481 TCCATCTTGAATGGGTCGCCCTGGCTGCCATPAGAACCGTGGCTGTGTGTGTCCAG 540
 DB 161 SerIlePheAspGlySerAlaLeuAlaValTrpIleValPheValValValGln 180
 QY 541 TACCGGCTAGGAATATTTGGTTTCTTCCACCATAGGATCAGCATGCTCCGGGAACTGG 600
 DB 181 TyrArgLeuGlnGlyIlePheGlyPhePheThrThrTrpAspGlnHisAlaProGlyAsnTrp 200
 QY 601 GCCTTCAGAGACCAAGTGGCTGCTGTCTGCTGCTGCTCCAGAAACATGAGTTCTTCGT 660
 DB 201 AlaPheValSerGlnValAlaAlaLeuSerTrpValGlnIleValSerIleGlnPhePheGly 220
 QY 661 GGGGACCCGAGCTGTGACCATCTTTGGGAGTCCGCGGAGCCATTAAGTTCTTCACT 720
 DB 221 GlyAspProSerSerValIleIlePheGlyGlySerAlaGlyAlaIleSerValSerSer 240
 QY 721 CTATACTGTCTCCATGAGCCAAAGGCTTATTCACAAAGCCATCATGAGAGTGGAGT 780
 DB 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHisIleValAlaIleMetGlySerGlyVal 260
 QY 781 GGCATCATCCCTTACCTGAGAGGCCCATGATTTAGAGAAAGTGAAGACTGCAAGTGT 840
 DB 261 AlaIleIleProTrpLeuGlnAlaHisAspTrpGlnLysSerIleValPheGlnValVal 280
 QY 841 GCACATTTCTGTGTGTAACAATGCGTCAAGACTGAGAGCCCTGTGAGGTCCTGAGAGAC 900
 DB 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlnIleLeuLeuArgCysLeuAlaGln 300
 QY 901 AAACCTCCAAAGAGCTGTGACCTTCAGCCAGAAAAACAAGCTTTCATCTGAGTGT 960
 DB 301 LysProSerIleValLeuLeuThrIleLeuSerGlnLysThrIleSerThrArgValVal 320
 QY 961 GATGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1020
 DB 321 AspGlyAlaPhePheProGlnGlnIleProLeuAspLeuLeuSerIleLysAlaPheLysAla 340
 QY 1021 ATTCTTCCATCATCTGAGATGATTAACAACAAGTGTGGCTTCTGCTGCTTGAAGAG 1080
 DB 341 IleProSerIleIleIleGlyValAlaAsnHisIleGlnCysGlyPheLeuLeuProMetLysGln 360
 QY 1081 GCTCTCGAAGATCTTCAGTGGCTTCAACAAGTCCCTTCCATCTGATTAACAAACATC 1140
 DB 361 AlaProGlnIleLeuSerIleLysSerAsnLysSerLeuAlaLeuHisIleLeuIleGlnAsnIle 380
 QY 1141 CTCGACATCCCGGCTCAGATATTGACACTGTGGCTTAAGATTAATCTTCATGACAAAGAC 1200
 DB 381 LeuHisIleProProGlnIleLysHisLeuValAlaAsnGlnTrpPheHisAspLysHis 400
 QY 1201 TCCCTGACTGAATCCGAGACAGTCTTCTGACTTGTGAGATGATGTCTTCTTGTGTGC 1260
 DB 401 SerLeuThrGlnIleIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValVal 420
 QY 1261 CCTGCACTGATCAACAGCTGATATACAGAGATGCTGGTGCACCTGTCTACTTCTATGAG 1320
 DB 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpGln 440
 QY 1321 TTTCCGACACCGGCTCAGTGTCTTGAAGACAGAAAGCCGCTTGTGTAAGCCGAGCAC 1380
 DB 441 PheArgHisArgProGlnCysPheGlyAspTrpLysProAlaPheValLysAlaAspHis 460
 QY 1381 GCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
 DB 461 AlaAspGlnValAlaArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
 QY 1441 GAAGAGCCAGGAGAGAGAGAAATTAATCTGAGCCGGAAGATGATGAATTAATCTGGCTCC 1500
 DB 481 GlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpAlaThr 500

QY 1501 TTGCTGGAACCGGGAATCTTAATGGGAACGACCTGCTCTGTGGCCAGCTTAATATCTG 1560
DB 501 PhalaehglntrgLyasnpProasngLyasnpaspleuserleutrrProaltrYasnmleu 520
QY 1561 ACTGAGCAGTACCTTCACAGCTGAGTGAACATGAGCCTCGGACAGAGACTCAAGAACCG 1620
DB 521 ThrGlnGlntrYleuGlnleuAspleuabMetSerleuGlnArgLeuYleuGlnPro 540
QY 1621 CGGCTGATTTTGTGACGACGACCATCCCTGATCTCTGTGCTCCGACATGCTCAC 1680
DB 541 ArgValasphetrPthSerThrleProleulleuSerAlasrAspMetleuHis 560
QY 1681 AGTCTCTTTCTCTCTTAACCTTCCTCTCTCTCCAGCCTTTCTTTTCTTTGAGCT 1740
DB 561 SerProleuserSerleutrrPheuserleuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
DB 581 Pro 581
RESULT 2
ADQ89094
ID ADQ89094 standard; protein; 581 AA.
XX
AC ADQ89094;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human urological disorder related protein 53010 SEQ:46.
XX
KM urological disorder; uropathic; cytostatic; urinary incontinence;
KM benign prostatic hyperplasia; human.
XX
OS Homo sapiens.
XX
PN WO2004065576-A2.
XX
PD 05-AUG-2004.
XX
PE 14-JAN-2004; 2004WO-US000750.
XX
PR 15-JAN-2003; 2003US-0440318P.
PR 04-FEB-2003; 2003US-0444783P.
PR 27-MAR-2003; 2003US-0457901P.
PR 08-MAY-2003; 2003US-0468775P.
PR 19-MAY-2003; 2003US-0471614P.
PR 16-JUN-2003; 2003US-0478742P.
PR 18-JUL-2003; 2003US-0488523P.
PR 30-JUL-2003; 2003US-0491156P.
PR 02-SEP-2003; 2003US-0495949P.
PR 26-SEP-2003; 2003US-0506332P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Karicheti V, Silos-Santiago I, Eliasof SD;
XX
XX WPI; 2004-562167/54.
DR N-PSDB; ADQ89093.
PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
PT 211 or for identifying a compound capable of treating a urological
PT disorder or identifying and treating a subject having a urological
PT disorder.
XX
XX
PS Claim 1; SEQ ID NO 46; 542pp; English.
XX
CC The present invention describes the use of polypeptides related to
CC urological disorders for identifying a compound capable of treating a
CC urological disorder, identifying a subject having a urological disorder,
CC or treating a subject having a urological disorder. Also described: (1) a
CC method for identifying a compound capable of treating a urological
CC disorder; (2) a method for identifying a subject having a urological

CC disorder; and (3) a method for treating a subject having a urological
CC disorder. The compound has uropathic and cytostatic activities. The
CC polypeptides related to urological disorders are useful for identifying a
CC compound capable of treating a urological disorder, identifying a subject
CC having a urological disorder, or treating a subject having a urological
CC disorder. Disorders include urinary incontinence and benign prostatic
CC hyperplasia. The present sequence represents a human urological disorder
CC related protein, which is used in the exemplification of the present
CC invention.
SQ Sequence 581 AA:
Alignment Scores:
Pred. No.: 3,216-287 Length: 581
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 97.13% Indels: 0
DB: Gaps: 0
US-10-023-515-3 (1-1746) x ADQ89094 (1-581)
QY 1 ATGCCACAGGAGCTTACTTCACTGCTTCAACATGTCCTTTTCTGATTCACGCC 60
DB 1 MetProGlnGlyLeuThrSerSerAlasrGlntrPcysPhePheLeuLleuGlnPro 20
QY 61 CTGTTGGGACACAGAGTGGGGAACCTGGGCTTCTGCTGAAGGCCACAGAGAAC 120
DB 21 LeuLeuGlnYlAsrGlntrPcysThrGlyProSerAlasrGlnGlyProGlnArgAsn 40
QY 121 ACCAGGCTGGATGATTCAGGAGCAAGTCACTGCTGCTGGAAGCCCTGCTCTG 180
DB 41 ThrArgLeuGlnYlPrlGlnGlnYlAsrGlnValThrValLeuGlnSerProValProVal 60
QY 181 AACGTGTTCTCGAGAGTCCCTTGTCTGCTCCCGCTGGAGATCCCTGCAATTAAGAAC 240
DB 61 AsnValPheLeuGlnYlAlaProPheAlaAlaProPheLeuGlnSerLeuArgPheThrAsn 80
QY 241 CCGAGAGCTGATGCGCTTGGATTAACCTTGGAGAACCTCTCCATTTGGGC 300
DB 81 ProGlnProAlasrProTrpAspAsnLeuArgGlnAlaThrSerYlProAsnLeuYs 100
QY 301 CTCGAGAACTCAGAGTGGCTCTCTTGAATCAACATGCTCAAGAGTCATTACCCGAA 360
DB 101 LeuGlnAsnSerGlntrPheLeuLeuAspGlnHisSheLeuYlValHisYlProYls 120
QY 361 TTGAGAGTGCAGAACTGCTCTTACCTGAACATCTAGTGGCTGCCACGCCGATACA 420
DB 121 PheGlyValSerGlnuAspCysLeuYlLeuAsnLleYlAlaProAlaHisAlaAspThr 140
QY 421 GGTTCGAAGCTCCCGCTTGGTGTGGTCCGAGAGAGTGGCTTCAGACCTGGCTCAGCC 480
DB 141 GlySerYlsLeuProValLeuValLtrPheProGlyGlyAlaPheYlSerGlySerAla 160
QY 481 TCCATCTTGAATGGAGTCCGCTGCTGCTGATGAGGAGCGTGTGTGATGTCAG 540
DB 161 SerLlePheAspGlySerAlaLeuAlaValYlGlnuAspValLeuValValGln 180
QY 541 TACCGGCTAGGAATATTTGTTCTTCAACATGAGATCGATCCCGGGAACGTG 600
DB 181 TyrArgLeuGlnYlPheGlnYlPhePheThrTrpAspGlnHisAlaProGlyAsntrP 200
QY 601 GCTTCAAGACCCAGGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 AlaPheYlsAspGlnValAlaAlaLeuSerTrpValGlnYsAsnLleGlnPhePheYl 220
QY 661 GGGAGCCCGAGCTGAGACCATCTTTGGCAGTCCGCGGAGGCATTAAGTTCAGT 720
DB 221 GlyAspProSerSerValThrLlePheGlyGlnSerAlaGlyAlaLleSerValSerSer 240
QY 721 CTATATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCCATCATGAGAGTGGGTG 780
DB 241 LeuLleuserPrometAlaYlGlyLeuPheHisYlAsnLleMetGlnSerGlyVal 260

```

OY 781 GCATCATCCCTTACCTGAGAGCCCATGATTATGAGAGAGTGGAGCCTTGAGAGTGGT 840
DB 261 AAlaIleIleProTyrIleuLulAhiAaSerYrGluLysSerGluuAapLeuGlnValVal 280
OY 841 GCACATTTCTGTGTGTAACAATGGTCAGACTTGAGAGCCCTGCTGAGTGGCTGAGAGCA 900
DB 281 AAlaIlePheCysGluYAsnAAsnAAsaSerAapSerGluuAlaLeuLysAkrGySLeuLkrGthr 300
OY 901 AAACCTCCAGAGAGTGTGACCCCTGACCCAGAAACAAGCTTTTCACTGAGTGGT 960
DB 301 LysProSerLysGluLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgValVal 320
OY 961 GATGAGTCTTTCTTCTATAGAGCTCTAGATCTTATGTTCTCAGAAAGCATTTAAAGA 1020
DB 321 AArgGluAlaPhePheProAAsnLulProLeuAAspLeuLeuSerGlnLysAlaPheLysAla 340
OY 1021 ATTCTTCATCATGATGAGAGTCAATTAACAAGAGTGGCTTCTGCTGCTATGAAGAG 1080
DB 341 IleProSerIleIleGluYAlaAAsnAAsnHsGluCysGluPheLeuLeuProMetLysGlu 360
OY 1081 GCTCCTGAGATCTCAGTGGCTCCAAAGTCCCTTGCCCTCCATGATCAAAACATC 1140
DB 361 AAlaProGluIleLeuSerGluSerAAsnLysSerLeuAlaLeuHsIleLeuIleGlnAAsnIle 380
OY 1141 CTGACATCCCGCCTCAGATATTGTGACACCTTGCTGCTAATGAATCTTCATGACAAGCAC 1200
DB 381 LeuHsIleProProGlnTyrIleuHsIleuValAlaAAsnGluTyrPheHsIAspLysHs 400
OY 1201 TCCCTGATGTAATCCGAGACAGAGTCTTGTGACCTTGCTGAGATGTTGTTGTGGTTC 1260
DB 401 SerLeuThrGlnLulLeuYAspSerLeuLysAAspLeuLysGluYAspValPhePheAlaVal 420
OY 1261 CCTGACATGATCAAGCTCGATATATCAAGAGATGCTGTGACACCTGTATCTTATGAG 1320
DB 421 ProAlaLeuIleThrAlaArgTyrHsIAspAAspAlaGluYAlaProValTyrPheLysGlu 440
OY 1321 TTTGGGACCGGAGCTCAGTGGCTTTGAAAGACAGAACCGGCTTTTGTCAAAACCGACAC 1380
DB 441 PheArgHsIAspProGlnCysPheGluuAAspThrLysProAlaPheValYAlaAspHs 460
OY 1381 GGTATGATGAAGTCCGCTTTGTGCTGGTGGTCCCTTCTGAAAGGGGACATTTATGTTCC 1440
DB 461 AAlaAspGluValAAspPheValPheGluYAlaPheLeuLysGluYAspIleValMetPhe 480
OY 1441 GAAGAGACCAAGAGAGAGAGAACTACTGAGCCGGAAGATGATGAATCTGGGCTACC 1500
DB 481 GluGluYAlaThrGluGluGluLysLeuLeuSerArgLysSheMetCysTyrTTPAlaThr 500
OY 1501 TTGGCTGGAACCGGGAATCTTAATGGAAACGACTGTCTCTGTGGCCAGCTTAAATCTG 1560
DB 501 PheAlaArgThrGluAAspProAAsnGluYAsnAAspLeuSerLeuTyrProLalaTyrAAsnLeu 520
OY 1561 ACTGAGAGAGTACCTCGAGCTGGAATTGAACATGAGCTTCGAGAGAGACATCAAGAACCG 1620
DB 521 ThrGluGlnTyrLeuGlnLeuAAspLeuAAsnMetSerLeuGlnAAspLeuLysGluPro 540
OY 1621 CGAGTGAATTTTGAACAGACCAATCCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
DB 541 ArgValAAspPheThrPheThrIleProLeuIleLeuSerHsIAsaSerAAspMetLeuHs 560
OY 1681 AGTCCCTTTCTTCTTAACCTTTCTCTCTCTCTCTCTGCAAGCCTTTCTTTTCTTTGTGCT 1740
DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
OY 1741 CCT 1743
DB 581 Pro 581

```

RESULT 3
ABP61004
ID ABP61004 standard; protein; 575 AA.
XX

```

AC ABP61004;
XX 10-SEP-2002 (First entry)
DT
XX
DE Novel human protein. SEQ ID 91.
XX
KW Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antitumor; vitruide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW and Alzheimer's disease; autoimmune disorder; parkinson's disease
KW inflammation; neoplastic diseases; nervous system related disorders and
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW haematological diseases; gastrointestinal ulceration; neuropathy;
KW disorders e.g. hypothyroidism; brain damages; colitis; cone photo-
KW transduction deficiency; neurological diseases; stroke; angiogenesis;
CC

```

The invention relates to an isolated polypeptide with signal sequences which allow it to be secreted extracellularly or membrane associated. The activity of polypeptides of the invention may be described as, CC cytosolic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic, CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory, CC cardiant, antitumor, vitruide, antithyroid, cerebroprotective, anorectic, CC and metabolic. Polypeptides and polynucleotides of the invention are CC useful in the treatment, or as a vaccine in the prevention of, cancer, CC wound healing disorders, infection, atherosclerosis, parkinson's disease CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder, CC inflammation, neoplastic diseases, nervous system related disorders and CC cardiovascular disorders; pancreatitis, respiratory disorder; CC hyperproliferation; systemic autoimmune disease; hyper-immunity, CC haematological diseases, gastrointestinal ulceration, neuropathy, CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo- CC transduction deficiency, neurological diseases, stroke, angiogenesis,

PI Agarwal P, Birkeland M, Cogswell JP, Kahnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
DR WPI; 2002-508784/54.
XX N-PSDB; ABQ86169.

PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 1(a); Page 312-313; 335pp; English.

CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosolic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antitumor, vitruide, antithyroid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders; pancreatitis, respiratory disorder;
CC hyperproliferation; systemic autoimmune disease; hyper-immunity,
CC haematological diseases, gastrointestinal ulceration, neuropathy,
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,

CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
CC growth abnormalities, and alleviation of precocious puberty. The
CC sequences given in records ABP60965-ABP61019 represent novel human
CC proteins of the invention

XX Sequence 575 AA;

Alignment Scores:

Pred. No.:	2,096-271	Length:	575
Score:	2915.00	Matches:	553
Percent Similarity:	97.54%	Conservative:	3
Best Local Similarity:	97.02%	Mismatches:	6
Query Match:	91.96%	Indels:	8
DB:	5	Gaps:	1

US-10-023-515-3 (1-1746) x ABP61004 (1-575)

```
QY 34 TGTGCTTTTTCGATTCCTCCAGCCCTGTTGGACACAGAGTGGGAAACTGGG 93
DB 14 TTPHAIETRPVALLAHLAHLAPro-----Thrylsely 25
QY 94 CTTTCGTGAAGGCGCCACAGAGAACACAGCGTGGATTCAGGCGAAGCATC 153
DB 26 ProserAlaGluGlyProGlnArgSantHrArgLeuGlyTTPILegLingLYlsGlnVal 45
QY 154 ACTGTGCGGGAAGCGCTGCTGTAACGTTTCCCTGGAGTCCCTTGTGCTGCTCC 213
DB 46 ThrvallLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaLaPro 65
QY 214 CCGGTGGGATCCCTGCGATTAACGAACCCGACGCTGCATGCGCCCTGGGATTAATTGCGA 273
DB 66 ProLeuGlySerLeuArgPheThrsanProGlnProAlaSerProTTPAspAsnLeuArg 85
QY 274 GAAGCCACTCTCACTAATTTGTGCTCAAGAACTGAGAGTGCCTCTTAGATCAA 333
DB 86 GluAlaThrSerTyrProAsnLeuGlyLeuGlnAsnSerGluTTPLeuLeuAspGln 105
QY 334 CACATGCTCAAGTGCATTACCCGAAATTCGAGAGTGAAGAACTGCGCTTACCTGAA 393
DB 106 HisMetLeuLysValHisTyrProLysPheGlyValSerGluAspCysLeuTyrLeuAsn 125
QY 394 ATCTATGCGCGTCCGACGCGGATCAAGGCTCAAGCTCCCGTCTTGATGTTCCCA 453
DB 126 IleTyrAlaProAlaHisAlaAspThrGlySerLysLeuProValLeuValITPphePro 145
QY 454 GAAGGTGCTTCAAGACTGCTCAGCTCAACTCTTGTGATGGGTCCGCTGCTGCTCAT 513
DB 146 GlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIaTyr 165
QY 514 GAGGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 573
DB 166 GluAspValLeuValValValValGlnTyrArgLeuGlyIlePheGlyPhePheThrThr 185
QY 574 TGGGATGACGATGCTCCGCGGGAACCTGAGCTTCAAGGACAGGAGGCTGCTGCTGCG 633
DB 186 TTPAspGlnHisAlaProGlyAsnTTPAlaPheLysAspGlnValAlaAlaLeuSerTTP 205
QY 634 GTCCAGAGAAACATGAGTTCCTTCGATGGGAGACCCAGCTCTGTGACCATCTTGGCGAG 693
DB 206 ValGlnLysAsnIleGluPhePheGlyGlyLysAspProSerSerValThrlIlePheGlyGlu 225
QY 694 TCCGCGGAGGACATTAAGTGTTCATAGTCTTATATCTGTCTCCATGCGCAAGGCTTATTC 753
DB 226 SerAlaGlyAlaIleSerValSerSerLeuLysLeuSerProMetAlaLysGlyLeuPhe 245
QY 754 CACAAAGGCATCATGAGAGTGGGTGGGCATCATCCCTTACCTGAGAGCCCATATTAT 813
DB 246 HisLysAlaAlaIleMetGlnSerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyr 265
QY 814 GAGAAGATGAGAGACTGCAAGTGTGTGCAATTTCTGTGTGAACATGCGTCAACTCT 873
DB 266 GluLysSerGluAspLeuGlnValValAlaHisPheCysGlyLysAsnAlaSerAspSer 285
```

```
QY 874 GAGGCCCTGTGAGTGTGCTGAGGACAAACCTTCAAGAGACTGTGACCCCTGAGCCAG 933
DB 286 GluAlaLeuLeuAspCysGlyLeuArgThrlYsPheSerLysGluLeuLeuThrlLeuSerGln 305
QY 934 AAAACAAAGTCTTTCACCTCGAGTGTGATGTGTCTTTCTTCTAATGAGCTTACAT 993
DB 306 LysThrLysSerPheThrArgValValAspGlyAlaPhePheProAsnGluProLeuAsp 325
QY 994 CTATTGCTCGAAGACATTTAAAGCAATTCCTCCATCATCGGAGTGCATTAACAGAG 1053
DB 326 LeuLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnAsnHisGln 345
QY 1054 TGTGAGCTTCCGTGCGCTGATGAAGAGGCTCTGAGATCTCACTGAGCTCCAAAGTCC 1113
DB 346 CysGlyPheLeuLeuProMetLysGluAlaProGluIleLeuSerGlySerAsnLysSer 365
QY 1114 CTTGCCCTCATCTGATACAAACATCTGACATCCCGCTCAGATATTGCACTTGTG 1173
DB 366 LeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleLeuProGlnTyrLeuHisIleVal 385
QY 1174 GCTAATGAATCTTCCATGAACAACCATCCCTGACTGAATCCGAGACAGCTTGTGAC 1233
DB 386 AlaAsnGluTyrPheHisAspLysHisSerLeuThrGluIleArgAspSerLeuLeuAsp 405
QY 1234 TTGCTTGAGATGATGTTCTTGTGAGTCCCTGCACTGATCAAGCTCGATATCAAGAGAT 1293
DB 406 LeuLeuGlyAspValPhePheValValProAlaLeuIleThrAlaArgTyrHisArgAsp 425
QY 1294 GCTGTGCACTGTCTACTTATGATGTTTCGGAACCGGCTCAGTGTCTTGAAGACAG 1353
DB 426 AlaGlyAlaProValTyrPheTyrGluPheArgHisArgProGlnCysPheGluAspThr 445
QY 1354 AAGCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTGTGAGTGTGCC 1413
DB 446 LysProAlaPheValLysAlaAspHisIleAlaAspGluValArgPheValPheGlyGlyAla 465
QY 1414 TTCCGAAAGGGGACATGTTATGTTTGAAGAGACAGAGAGAGAGAAATTACTGAGC 1473
DB 466 PheLeuLysGlyAspIleValMetPheGluGlyAlaThrGluGluGluLysLeuSer 485
QY 1474 CGAAGATGATGAATAATACCTGGGCTACCTTGTCTGGAACCGGAAATCTTAATGGAAAGAC 1533
DB 486 ArgLysMetMetLysTyrTTPAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 505
QY 1534 CTGTCTGTGTGGCCAGCTTAATATCTGACTGAGACAGTACTCTCAAGCTGAACATG 1593
DB 506 LeuSerLeuTTPProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMet 525
QY 1594 AGCCTCGGACAGAGACTCAAGAAACCGGAGTGAATTTTGAACAGACCATCCCGCTG 1653
DB 526 SerLeuGlyGlnAspGlyLysGluProArgValAspPheThrPheThrIlePheProLeu 545
QY 1654 ATCTGTCTGCTCGACATGCTCCACAGTCTCTTCTTCTTAACCTTCTCTCTCTC 1713
DB 546 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 565
QY 1714 CTCAGCTTCTTCTTCTTCTTGTGCTCT 1743
DB 566 LeuGlnProPhePhePheCysAlaPro 575
```

RESULT 4
ID AAE25025 standard; protein; 642 AA.
XX AAE25025;
AC AAE25025;
DT 30-OCT-2002 (first entry)
XX Human drug metabolising enzyme (DME-10).
XX Human; drug metabolising enzyme; autoimmune; inflammatory disorder;
XX Human; drug metabolising enzyme; autoimmune; inflammatory disorder;
KW acquired immunodeficiency syndrome; AIDS; atherosclerosis; psoriasis;

KM proliferative disorder; arteriosclerosis; cirrhosis; hepatitis; cancer;
KM aetmia; neurological disorder; Alzheimer's disease; Huntington's disease;
KM dementia; Parkinson's disease; developmental disorder; anaemia; adenoma;
KM drug screening; endocrine disorder; conjunctivitis; glaucoma; cataract;
KM renal tubular acidosis; eye disorder; epilepsy; thrombosis; peptic ulcer;
KM anorexia; metabolic disorder; cystic fibrosis; diabetes; liver disorder;
KM goitre; gastrointestinal disorder; gene therapy; viroicide; anticoagulant;
KM anticonvulsant; noctropic; enzyme; DME-10.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..50
FT /label= Signal_peptide
FT Domain 32..56
FT /note= "Transmembrane domain"
FT Protein 51..642
FT /note= "Mature human DME-10"
FT Domain 113..135
FT /note= "Transmembrane domain"
FT Domain 204..220
FT /note= "Transmembrane domain"
FT Domain 234..250
FT /note= "Transmembrane domain"
FT Domain 287..314
FT /note= "Transmembrane domain"
FT Domain 463..491
FT /note= "Transmembrane domain"
XX
PN WO200246426-A2.
XX
PD 13-JUN-2002.
XX
PF 04-DEC-2001; 2001WO-US047429.
XX
PR 08-DEC-2000; 2000US-0254308P.
PR 15-DEC-2000; 2000US-0256189P.
PR 21-DEC-2000; 2000US-0257713P.
PR 19-JAN-2001; 2001US-0262706P.
PR 02-FEB-2001; 2001US-0266020P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Sanjanwala MM, Yao MG, Au-Young J, Baughn MR, Arvizu C, Ring HZ,
PI Lee EA, Ding L, Hafalia AJA, Tang YT, Yue H, Tribouley CM, Lu DM,
PI Lal PG, Warren BA, Yang J, Walla NK, Nguyen DB, Gandhi AR, Lu Y,
PI Ison CH;
XX
DR WPI: 2002-519668/55.
XX N-PSDB: AAD40574.
XX
PT Novel human drug metabolizing polypeptide, useful in diagnosis,
PT prevention or treatment of autoimmune/inflammatory, cell proliferative,
PT neurological, developmental, endocrine, metabolic and gastrointestinal
PT disorders.
XX
PS Claim 65; Page 155-156; 169pp; English.
XX
CC The invention relates to an isolated human drug metabolizing enzyme (DME)
CC and its nucleotide. DME is useful for diagnosing, treating or preventing
CC disorders associated with aberrant expression of DME, where the disorders
CC are selected from autoimmune/inflammatory disorder such as acquired
CC immunodeficiency syndrome (AIDS), asthma, arteriosclerosis, psoriasis,
CC uveitis; a cell proliferative disorder such as arteriosclerosis,
CC cirrhosis, hepatitis, and cancer; a neurological disorder such as
CC Alzheimer's disease, Huntington's disease, dementia, Parkinson's disease;
CC a developmental disorder such as renal tubular acidosis, epilepsy,
CC anaemia; an endocrine disorder such as adenoma, thrombosis and infections
CC ; an eye disorder such as conjunctivitis, glaucoma, cataract; metabolic
CC disorder such as cystic fibrosis, diabetes and goitre; a gastrointestinal
CC disorder such as anorexia, peptic ulcer; and liver disorders. DME is
CC useful in a number of drug screening techniques and to analyse the
CC proteome of a tissue or cell type. The invention is useful for creating

CC knock-in humanised animals or transgenic animals to model human diseases,
CC in somatic or germline gene therapy, to generate a transcript image of a
CC tissue or cell type, for detecting differences in the chromosomal
CC location due to translocation, inversion, etc. among normal, carrier or
CC affected individuals, and as hybridisation probes for mapping naturally
CC occurring genomic sequences. The present sequence is human DME-10
XX
SQ Sequence 642 AA;

Alignment Scores:

Pred. No.:	6,21e-270	Length:	642
Score:	2900.00	Matches:	548
Percent Similarity:	99.82%	Conservative:	2
Best Local Similarity:	99.46%	Mismatches:	1
Query Match:	91.48%	Indels:	0
DB:	5	Gaps:	0

US-10-023-515-3 (1-1746) x AAE25025 (1-642)

QY	91	GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAAGGCCAAGCAA	150
DB	92	GLYProSerAlaIugIyProGlnArgAsnThrArgLeuGlyTrpIleGlnGlyLysGln	111
QY	151	GTCACGTGCTGGGGAAGCCCTGTCGTCGAGAGTGTCTCGAGATCCCTTGTGCTGCT	210
DB	112	ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAla	131
QY	211	CCCCGCTGGGATCCCTGCGATTACGAACCCGACGCTGCATCGCCCTGGATTAATTG	270
DB	132	ProProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTrpPheAsnLeu	151
QY	271	CGAAGAGCCACCTCCACCTTAATTTGTGCTCAGAACTCAGAGTGGCTCTTAGAT	330
DB	152	ArgGlnIaThrSerxerYrProAsnLeuCyLeuGlnAsnSerGlyTrpLeuLeuAsp	171
QY	331	CAACACATGCTCAAGGTGATTAACCCGAATTGAGATGTCAGAAAGCTGCCCTTACCTG	390
DB	172	GlnHisMetLeuLysValHisTyrTrpPheLysPheGlyValSerGlnAspCyLeuTrpLeu	191
QY	391	AACATCTATGCGCTGCGCCACCGCCGATACAGGCTCCAGCTCCCGCTTGTGTGCTTC	450
DB	192	AsnIleTyrAlaProAlaHisAlaAspThrGlySerLeuProValLeuValTrpPhe	211
QY	451	CCAGAGAGTGGCTTCAAGCTGCTCAGCTCCATCTTGAATGGGTCCGCTGCTGCTGCC	510
DB	212	ProGlyGlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaAla	231
QY	511	TATGAGACGTGCTGTGTGCTGCTGCATACCGGCTAGAAATATTGGTTCTTACACC	570
DB	232	TyrGlnAspValLeuValValValValGlnTyrArgLeuGlyIlePheGlyPhePheThr	251
QY	571	AACATGGAATCAGACATGCTCCGGGAACTGGGCTTCAAGAACACAGATGGCTGCTGCC	630
DB	252	ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSer	271
QY	631	TGGGTCAGAAAGACATCGATTCTTCGGTGGGAGCCGAGCTCGTACCATCTTGGC	690
DB	272	TrpValGlnLysAsnIleGlnPheGlyValAspProSerSerValThrIlePheGly	291
QY	691	GAGTCCGCGGAGGCATATAGTGTTCATGCTTATACGTCTCCATGCGCAAGGCTTA	750
DB	292	GlnSerAlaGlyAlaIleSerValSerSerLeuSerPheMetAlaLysGlyLeu	311
QY	751	TTCCCAAAAGCATATGAGAGTGGGTGGCCATCATCCCTTACTCTGAGGCCCATAT	810
DB	312	PheHisLysAlaAlaIleMetGlnSerGlyValAlaIleIleProTyrLeuGlnAlaHisAsp	331
QY	811	TATGAGAAAGATGAGACCTCAGAGTGTGTCACATTTCTGAGGAACAATGGCGAGAC	870
DB	332	TyrGlnLysSerGlnAspLeuGlnValValAlaHisPheCysGlyAsnAsnHisSerAsp	351
QY	871	TCTGAGCCCTGCTGAGGTGCTTGAAGCAAAACCTTCGAAGAGCTGACCTTCAGC	930


```

Db      352 SerGluAlaLeuLeuArgCysLeuArgThrIysProSerIysGluLeuLeuThrLeuSer 371
Qy      931 CAGAAAAGAAAGCTTTTACCTGAGTGTGATGCTGCTTTCTTCCATAGAGCTCTTA 990
Db      372 GlnIysThrIysSerPheThrArgValAlaIspGlyAlaPhePheProAsnGluProLeu 391
Qy      991 GATCTATTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAC 1050
Db      392 AspLeuLeuSerGlnIysAlaPheIysAlaIleProSerIleIleGlyValAlaAsnHis 411
Qy      1051 GAGGTGGCTTCTGCTGCTCCATGAGAGAGGCTCTGAGATCTCAGTGGCTCCACAG 1110
Db      412 GluCysGlyPheLeuLeuPheProMetIysGlyAlaProGluValLeuSerGlySerAsnIys 431
Qy      1111 TCCCTTGGCCCTCCATCGATCAAAAGCAATCCTGCAATCCGCGCCATATTTGACCTT 1170
Db      432 SerLeuAlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIyIysLeuHis 451
Qy      1171 GTGGCTAATGAATCTTTCATGACAGACATCCCTGACCTGAATCCAGACAGTCTTCTG 1230
Db      452 ValAlaAsnGlnIyIysPheHisAspIysHisSerLeuThrGlnIleArgAspSerLeuLeu 471
Qy      1231 GACTTGTCTGAGAGATGTGTCTTGTGTGCTCCCTGACATGATCAGCTCGATATACAGA 1290
Db      472 AspLeuLeuGlnIysArgValAlaPheValAlaProAlaLeuIleThrAlaArgIyHisArg 491
Qy      1291 GATGCTGGTGAACCTGTCTACTTCTATGAGTTTGCGGACCGGCGCTCAGTCTTGAACAC 1350
Db      492 AspAlaGlyAlaProValIyIysPheIyIysGlnPheArgHisAspProGlnIyIysPheGlnAsp 511
Qy      1351 ACAGAGCGGCTTTTGTCAAGCCGACCAAGCTGATGAGTCCGCTTGTGTGCGTGTG 1410
Db      512 ThrIysProAlaPheValIyIysAlaAspHisAlaAspGluValAlaArgPheValIyIysGly 531
Qy      1411 GCCTTCTGAAAGGCGGACATTTGATGTTTGAAGAGCCACGAGAGAGAGAACTTACTG 1470
Db      532 AlaPheLeuIyIysGlyAlaPheIleValMetPheGlnIyAlaThrGlnIyGlnIyIysLeuLeu 551
Qy      1471 AGCGGGAAGATGATGAATAATCTGGGCTACCTTGTCTGCAACCGGGAATCCCTATGGGAAAC 1530
Db      552 SerIyIysMetIyIysIyIysIyIysAlaThrPheAlaAspGlnIyIysAsnProAsnIyIysAsn 571
Qy      1531 GACCTGTCTGTGGGCGGCTTATATATCTGACTGAGTCACTCCAGCTGGAATCTTGTGAC 1590
Db      572 AspLeuLeuLeuIyIysProAlaIyIysAsnLeuThrGlnIyIysIyIysLeuIyIysAsn 591
Qy      1591 ATGAGCCTCGGACAGAGCTCAAAAGAACCGGCGGTGATTTTGGACGACCAATCCCC 1650
Db      592 MetSerLeuGlnIyIysAlaArgLeuIyIysGlnIyIysProAlaArgValGlnPheIyIysThrSerThrIlePro 611
Qy      1651 CTGATCTGTGTGCTGCTCGGACATGCTCCAGAGTCTCTTCTTCTTAACTTCTCTCT 1710
Db      612 LeuIleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSer 631
Qy      1711 CTCCCTCAGAGCTTTCTTTTCTTTTGTGCTCCT 1743
Db      632 LeuLeuGlnIyIysPheIyIysPheIyIysPheIyIysAlaPro 642

```

```

XX      25-MAR-2004.
PD      12-SEP-2003; 2003WO-US028227.
XX      12-SEP-2002; 2002US-0410259P.
XX      12-SEP-2002; 2002US-0410259P.
XX      12-SEP-2002; 2002US-0410260P.
XX      (INCY-) INCYTE CORP.
XX      Schmidt JP, Wright RJ, Bruns CM, Marianovic MM, Shen F,
PI      Hartschorn TA, Suchorski MT, Altus CM, Plets SO, Elder LV,
PI      Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP,
PI      Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH,
PI      Petralla CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LU,
PI      Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES,
PI      Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D,
PI      Fectury S, Shi X, Suarez CJ,
XX      MPI; 2004-329368/30.
DR      N-PsDB; ACN42766.
XX      New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT      in diagnosing a condition, disease or disorder associated with human
PT      molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT      in gene mapping.
XX      Claim 27, Page; 190pp; English.
XX      The invention relates to novel diagnostic and therapeutic polynucleotides
CC      selected from one of the 2722 sequences defined in the specification. A
CC      polynucleotide of the invention may have a use in gene therapy. The human
CC      diagnostic and therapeutic polynucleotides (dthp) or polypeptides associated
CC      used to diagnose a particular condition, disease or disorder associated
CC      with human molecules, e.g. cell proliferative disorders,
CC      autoimmune/inflammatory disorders, developmental disorders, endocrine
CC      disorder, neurological disorders, gastrointestinal disorders, or
CC      infections caused by virus, bacteria, fungi or parasites. The dthp
CC      molecules may also be used in genetic mapping, in identifying individuals
CC      from minute biological samples, in detecting single nucleotide
CC      polymorphisms, as molecular weight markers, and for somatic or germline
CC      gene therapy. The present sequence represents a dthp protein of the
CC      invention. Note: The sequence data for this patent is not represented in
CC      the printed specification, but was obtained in electronic format directly
CC      from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX      SQ      Sequence 642 AA;
XX      Alignment Scores:
XX      Pred. No.:      6,21e-270      Length:      642
XX      Score:      2900.00      Matches:      549
XX      Percent Similarity:      99.82%      Conservative:      1
XX      Best Local Similarity:      99.64%      Mismatches:      1
XX      Query Match:      91.48%      Indels:      0
XX      DB:      8      Gaps:      0
XX      US-10-023-515-3 (1-1746) x ABM84114 (1-642)
Qy      91 GGCGCTTCTGTGTAAGGGCCACAGAGGAACACAGCTGGATGATTCAGGGCAAGCA 150
Db      92 GlyProSerAlaGlnIyIysProGlnArgAsnThrArgLeuGlnIyIysIleGlnIyIysGln 111
Qy      151 GTCACTGTGTGGGAGAGCCCTGTGCTGTGAAGTGTCTTCTGGAGTCCCTTTGCTGCT 210
Db      112 ValThrValLeuGlnIyIysSerProValAlaAsnValPheLeuGlnIyIysAlaIle 131
Qy      211 CCCCCTGGATGATCCCTGCGATTTAAGAACCCGAGAGCTGATGAGCTGGGATTAATCTTG 270
Db      132 ProProLeuGlnIyIysSerLeuArgPheThrAsnProGlnProAlaSerProIyIysPheAsnLeu 151
Qy      271 CAGAGGCAACCTCTACCTTAATTTGTGCTTCAGAACTCAGAGTGGCTGCTTATAG 330
Db      152 ArgGlnAlaThrSerIyIysProAsnLeuIyIysGlnAsnSerGlnIyIysPheLeuLeuAsp 171

```

```

RESULT 5
ABM84114
ID      ABM84114 standard; protein; 642 AA.
XX
AC      ABM84114;
XX
DT      18-NOV-2004 (first entry)
XX
DE      Human diagnostic and therapeutic protein SEQ ID NO:4363.
XX
KM      gene therapy; human diagnostic and therapeutic polynucleotide; dthp.
XX
OS      Homo sapiens.
XX
PN      WO2004023973-A2.

```

```

QY 331 CAACATGCTCTCAAGGTGATACCCGAAATTGAGTGTGAGAGACTGCTTACCTG 390
DB 172 GlnHISMetLeuLysValIhSTyRProLysPheGlyValSerClnuSPryLeuLysLeu 191
QY 391 AACATGTATGCGCTGCGCCACCGCGATACAGGCTCCAGCTCCCGCTTGTGTGGTTC 450
DB 192 AamIleTyAlaProAlaIhValaAerThnGlySerLysLeuProValLeuValTlPhe 211
QY 451 CAGAGAGGTGCTTCAAGACTGCTCAGCTCCATCTTTGATGGGTCCCGCTGGCTGCC 510
DB 212 ProGlyGlyAlaPheLysThnGlySerAlaSerTllePheASPGLySerAlaLeuAla 231
QY 511 TATGAGACGTGCTGTTGTGCTGCTCAGTACCGGCTAGAAATTTTGGTTCTTCC 570
DB 232 TyRgluAerValLeuValValValGlnTyAlaGleuGlyLlePheGlyPhe***Ser 251
QY 571 ACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGCTGCTCC 630
DB 252 ThrtTPAerGlnHISAlaProGlyAanTlPAlaPheLysAerPGLInValAlaAlaLeuSer 271
QY 631 TGGGTCCAGAAAGATGAGTCTTCCGTGGGGAACCCAGCTCTGATCCATCTTTGGC 690
DB 272 TlPValGlnLysAanIleGlnPheGlyGlyAerProSerSerValThrtllePheGly 291
QY 691 GAGTCCGCGGAGCCATAGTGTCTTCTAGTCTTATATGCTCTCCATGGCCAAAGCTTA 750
DB 292 GluSerAlaGlyAlaIleSerValSerLeuIleLeuSerProMetAlaLysGlyLeu 311
QY 751 TTCCCAAAAGCCATGATGAGAGTGGGTGGGATCATCCCTTACCTGAGAGCCATGAT 810
DB 312 PheHISLysValAlleMetGluSerGlyValAlaIleIleProTyLleuGlnAlaIhAerP 331
QY 811 TATGAGAAAGTGAAGACTGTCAGGTGGTGGACATTTCTGTGTGAATAGCTGCAGAC 870
DB 332 TyRgluLysSerGluAerPheGlnValValaIhAerPheCyGlyAanAanIleAerP 351
QY 871 TCTGAGGCGCTGCTGAGGTGCTGAGGACAAACCCCTCAAGAGACTGTCAGCCCTGAC 930
DB 352 SerGlnAlaLeuLeuArgCyLysLeuArgThrtLysProSerLysGlnLeuLeuThlLeuSer 371
QY 931 CAGAAACAAAGTCTTTCACCTCGAGTGTGATGCTGCTTCTTCCAAAGACCTCTA 990
DB 372 GlnLysThrtLysSerPheThnArgValaAerPGLyAlaPhePheProAerGlnProLeu 391
QY 991 GATGTATTTGCTCAGAAAGCATTTAAAGCAATCTTCCATCATCGAGTGAATTAACAC 1050
DB 392 AarPheLeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValaAanAanHIS 411
QY 1051 GAGTGTGGCTTCCGTGCTGCTGATGAAGAGGCTCTGAGATCTCGAGTGGCTCAACAG 1110
DB 412 GluCyAerGlyPheLeuLysProMetLysGlnAlaProGlnIleLeuSerGlySerAanLys 431
QY 1111 TCCCTTGCCCTCCATGATGATCAAAACACTCTGACACTCCCGGCTCGATTTTGGACCT 1170
DB 432 SerLeuAlaLeuHISLeuIleGlnAanIleLeuHISIleProProGlnLtyLleuHISLeu 451
QY 1171 GTGGCTATATGATATCTTCAAGACACTCCCTGATCTGAATCCGAGACAGTCTTGTG 1230
DB 452 ValAlaAanGlnLtyRPhenIleAerLysAerLysSerLeuThnGlnIleArgAerSerLeuLeu 471
QY 1231 GACTGTGTTGAGATGTGTTCTTGTGTGCTCCGCTGATGATCAAGTGCATATCAGAGA 1290
DB 472 AarPheLeuLeuLysAerValPhePheValaProAlaLeuIleThnAlaArgTyRHSArg 491
QY 1291 GATGCTGCTGACACTGTCTATCTTATGATTTCCGACCGGCTCAGTCTTGTGAAGAC 1350
DB 492 AarPAlaGlyAlaProValTyRPhetyRgluPheArgHISArgProGlnLysPheGlnAerP 511
QY 1351 ACAGAGCGCGCTTTGTCAAGCGACACAGCTGATGAAGTCCGTTGTTCGCTGGT 1410
DB 512 ThrtLysProAlaPheValLysValaAerPAlaIleAerPGLuValArgPheValPheGlyGly 531

```

```

QY 1411 GCCTTCTGAAGGGGAGATTTGTTATGTGCAAGAGCCACGAGAGAGAGATTACTG 1470
DB 532 AlaPheLeuLysGlyAerPAlaIleValMetPheGlnGlyAlaThnGlnGlnGlnLysLeu 551
QY 1471 AGCCGGAAGATGAAGAAATTACTGGGCTACCTTTGCTGGAACCGGGGAATCCATAAGGGAGAC 1530
DB 552 SerArgLysMetMetLysTyRTPAlaThnPhenAlaArgThnGlnLysAanPheGlnLys 571
QY 1531 GACTGTCTCTGTGGCCAGCTTAATATGACTGAGCAGTACCTTCAGCTGAGACTTGAAC 1590
DB 572 AarPheLeuSerLeuTlPProAlaTyRAsnLeuThnGlnGlnTyRLeuGlnLeuAerPLeuAan 591
QY 1591 ATGAGCCTCCGACAGAGACTCAAGAAACCGGGGGTGAATTTTGGACACAGACCATGCC 1650
DB 592 MetSerLeuGlnGlnArgLeuLysGlnProArgValaAerPheTlPnSerThrtllePro 611
QY 1651 CTGATCTGTCTGCTGCTCCGACATGCTCCACAGTCCCTCTTCTTCTTCAACTTCCCTCT 1710
DB 612 LeuIleLeuSerAlaSerAerPheMetLeuHISerProLeuSerSerLeuThnPhenLeuSer 631
QY 1711 CTCTCCAGGCTTTCTTTTCTTGTGCTGCT 1743
DB 632 LeuLeuGlnProPhePhePhePheCyAlaPro 642
RESULT 6
ABP61005
ID ABP61005 standard; protein; 581 AA.
AC ABP61005;
XX 10-SEP-2002 (first entry)
DT 10-SEP-2002
XX
DE Novel human protein. SEQ ID 92.
XX
KW Human; cytosolic; vulnary; antiarteriosclerotic; antiparkinsonian;
KW nootropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antidiabetic; antidiabetic;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW angioneuema; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty.
OS Homo sapiens.
XX
PN NC0200250105-A1.
XX
PD 27-JUN-2002.
XX
PF 17-DEC-2001; 2001WC-US049232.
XX
PR 19-DEC-2000; 2000US-0256710P.
PR 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
PA (SMK ) SMITHKLINE BEECHAM CORP.
PA (SMK ) SMITHKLINE BEECHAM PLC.
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y,

```



```

OY 1174 GCTATGAATATTCTTCATGACAAGCACTCCGATCTGAAATCCGAGACAGTCTTGGAC 1233
DB 386 AAlaenGluTYrPheHlaAspYehHsSerLeuThrGluLLeuAspSerLeuLeuAsp 405
OY 1234 TTGCTGGAGATGCTGTTCTTTGTGCTGCTGCTGATCAAGCTGATATCACAGAGAT 1293
DB 406 LeuLeuGluYalPheValPhePheValProAlaLeuLLeuThrAlaGlyTyrHisArg--- 424
OY 1294 GCTGTGACCTGCTTACTTCTATGAGTTTGCGACCGCGCTCACTGCTTGAACAACAG 1353
DB 424 ----- 424
OY 1354 AAGCGGCTTTTTCGAAGCCGACACCGCTATGAAAGTCGCTTGTGTCGTGTGCC 1413
DB 424 ----- 424
OY 1414 TTCCTGAAGGGGACATTTGTAATGTTTCGAAGAGACCGACGAGAGAGATTACTGAGC 1473
DB 425 -----GluGlyAlaThrGluGluGluLeuLeuSer 435
OY 1474 CGGAAGATGATGAATATCTGGGCTACCTTTGCTGCAACCGGGAATCTTAATGGAGACAC 1533
DB 436 ArgGlySerMetCysTyrTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAsp 455
OY 1534 CTGTCTCTGTCGCGACCTTATATCTGACTGACGATCACTCCGCTGCTGATGAACATG 1593
DB 456 LeuSerLeuThrProAlaTyrAsnLeuThrGluGlnTyrLeuGlnLeuAspLeuAsnMet 475
OY 1594 AGCTTCGACAGACAGCTCAAAAGAACCGGCGGTGATTTTGGACGACGACCATCCCGCTG 1653
DB 476 SerLeuGluGlnArgLeuLeuGluProArgValaIaPheThrSerThrIleProLeu 495
OY 1654 ATCTGCTGCTCTCCGACATGCTCAAGTCTCTTTCTTCTTAATCTTCTCTCTC 1713
DB 496 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeu 515
OY 1714 CTCACGCTTCTTTCTTTTGTGCTCCT 1743
DB 516 LeuGlnProPhePhePheCysAlaPro 525

RESULT 8
ABUS4639 ID ABUS4639 standard; protein; 581 AA.
AC ABUS4639;
XX 03-JUN-2003 (first entry)
DT
PT
XX Human NOXV polypeptide #98.
DE
XX Human; NOXV; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
OS Homo sapiens.
XX
XX WO200281498-A2.
XX 17-OCT-2002.
XX
XX 03-APR-2002; 2002WO-US010780.
XX
XX 03-APR-2001; 2001US-0281086P.
XX 03-APR-2001; 2001US-0281136P.
XX 05-APR-2001; 2001US-0281863P.
XX 05-APR-2001; 2001US-0281906P.
XX 06-APR-2001; 2001US-0282020P.

```

```

PR 10-APR-2001; 2001US-0282930P.
PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285899P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.

(CDRA-) CDRAGEN CORP.
XX
XX Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA,
PI Patuxian M, Liu X, Gusev VY, Li L, Vernet CM, Zethusen BD,
PI Gorman L, Shenoy SG, Pena CE, Smithson G, Burgess CE, Gerlach V,
PI Padigaru M, Shinkels RA, Gangoli EA, Taupier RJ, Caeman SJ, Ji W,
PI Anderson DM, Leite MW, Rastelli L, Edinger SR, Stone DJ,
PI MacDougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA,
PI Ellerman K,
XX
XX MPI: 2003-046858/04.
XX N-P8DB; ABX72267.
XX
XX New isolated NOXV polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
XX Claim 1; Page 303; 666pp; English.
XX
XX The invention relates to human polypeptides, termed NOXV, and the
CC polymucleotides encoding them. The polypeptides and polymucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious diseases, Crohn's disease, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABUS4542-ABUS4647 represent human NOXV polypeptides
CC of the invention
XX
XX SQ Sequence 581 AA;

```

Alignment Scores: 8.79e-241 Length: 581
Pred. No.: 2597.50 Matches: 497
Percent Similarity: 95.07% Conservative: 4
Best Local Similarity: 94.31% Mismatches: 17

Query Match: 81.94% Indels: 9
 DB: 6 Gaps: 3
 US-10-023-515-3 (1-1746) x ABUS4639 (1-581)

```

QY 91 GGGGCTTGTCTGAAGGGCCACAGAGAACACACAGGCTGGATGATTCAGGGCAAGCA 150
DB 27 GTPROSERAlAGLUGlYProGlnrGaenThrArgLeuGlyTrrlGelnGlyysGln 46
QY 151 GTCATCTGTCTGGAAAGCCCTGTGCTGGAACGTGTCTCTGGAGTCCCTTGTGCT 210
DB 47 ValThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaIa 66
QY 211 CCCCCGTGGGATCCCTGGATTTTGAACCCGAGCCCTGCATGCCCTGGGATTAATCTG 270
DB 67 ProProLeuGlySerLeuThrPheThrAsnProGlnProAlaSerProTrpAsnLeu 86
QY 271 CGAAGAGCCACCTCTTACCTTAATTTGTGCTTCAAGACTCAGAGTGGCTGCTTAGAT 330
DB 87 ArgGlnAlaThrSerTrpProAsnLeuCySLeuGlnAsnSerGlnTrpLeuLeuAsp 106
QY 331 CAACGATGCTCAAGTGAATTAACCCGAATTCGAGAGTGTCAAGAACTGCTCTTACCTG 390
DB 107 GlnHisMetLeuYsValHisIeTrpProYsPheGlyValSerGlnAspCySLeuYzLeu 126
QY 391 AACATCTATGGCGCTGCCACGCGCATACAGGCTCAGAGCTCCCGCTTGGTGTGCTTC 450
DB 127 AsnIeTrpAlaProAlaHisAlaAspThrGlySerYsLeuProValLeuValTrpPhe 146
QY 451 CCAGAGGCTGCTTCAAGACTGCTGACCTTCATCTTTGATGGTCCGCTGGCTGCC 510
DB 147 ProGlyGlyAlaPheYsThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIa 166
QY 511 TATGAGACGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 570
DB 167 TyrGlnAspValLeuValValValValGlnTrpHisLeuGlyIlePheGlyPhePheThr 186
QY 571 ACATGGGATCAGATGCTCCGCGGAACTGGGCTTTCAGAGACAGAGTGGCTGTCTGCC 630
DB 187 ThrTrpAspGlnHisAlaProGlyAsnTrpAlaPheYsAspGlnValAlaHisLeuSer 206
QY 631 TGGGTCCAGAAAGAACATGAGTTCCTCGGTGGGACCCACAGCTCTGTGACATTTTGGC 690
DB 207 TrpValGlnYsAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGly 226
QY 691 GAGTCCGCGGAGGCATTAAGTGTCTTACTTATGCTGTCTCCATGGCCAAAGGCTTA 750
DB 227 GluSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaYsGlyLeu 246
QY 751 TTCCCAAAAGCCATCATGAGAGAGTGGGATGCGCATCATCCCTTACCTGAGGCCATGAT 810
DB 247 PheHisIeYsAlaIleMetGluSerGlyValAlaIleIleProYrLeuGlnAlaHisAsp 266
QY 811 TATGAGAAAGTGAAGACCTGACAGTGTGTGCAATTTCTGTGTAAACAATGCTGTGAC 870
DB 267 TyrGlnYsSerGlnAspLeuGlnValValAlaHisPheCySgIyAsnAsnAlaSerAsp 286
QY 871 TCTGAGGCGCTGTGAGTGTCTGAGAGCAAAACCTTCAAGAGAGTGTGACCTGAGC 930
DB 287 SerGlnAlaLeuLeuAspGlySLeuAspArgThrIysProSerYsGlnLeuLeuThrLeuSer 306
QY 931 CAGAAAAGAAAGTCTTCACTCGAGTGTGTGATGTGCTTTCTTCTTAATGAGCCCTTA 990
DB 307 GlnIysThrYsSerPheThrArgValValAspGlyAlaPhePheProAsnGlnProLeu 326
QY 991 GATCTATGTCTGAGAAAGATTAAAGCAATCTTCCATCATCCGAGAGCATTAACAC 1050
DB 327 AspLeuLeuSerGlnYsAlaPheYsAlaIleProSerIleIleGlyValAlaAsnHis 346
QY 1051 GAGTGTGGCTTCTGTGCTGCTATGAGAGAG-----GCTCTGAGATCCAGTGGCTCC 1104
DB 347 GluCyGdlYpHeuLeuLeuProMetValArgIleLeuAlaValHisThrAlaThrProSer 366

```

```

QY 1105 AACAGATCCCTTGGCCCTC-----CATCTGATACAAACATCTTGCAATC 1149
DB 367 AsnArgAspAlaValAlaLeuAlaSerThrAlaGlyHisPheHisArgArg-----HisIle 384
QY 1150 CCGGCTCAGATTTTGGACCTTGTGGCTTAATGATATCTCCATGACAGACATCCCTGAT 1209
DB 385 ProGlnGlnTrpLeuHisSLeuValAlaAsnGlnTrpPheHisAspYsHisSerLeuThr 404
QY 1210 GAAATCCGAGACAGTCTTCTGAGCTTGTGAGATGTGTCTTGTGTGCTCTGCACTG 1269
DB 405 GlnIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValValProAlaLeu 424
QY 1270 ATCAGAGTGTATTCACAGAGATGCTGTGCACCTGTCTTACTTATGAGTTTGGCAC 1329
DB 425 IleThrAlaArgYrHisArgAspAlaGlyAlaProValYrPheYrGlnPheArgHis 444
QY 1330 CGGCTCAGTGTCTTGAAGACACAGACAGCCGCTTTGTCAAGCCGACAGCTGATGAA 1389
DB 445 ArgProGlnCySProGlnAspThrIysProAlaPheValYsAlaAspHisAlaAspGln 464
QY 1390 GTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1449
DB 465 ValArgPheValPheGlyGlyAlaPheLeuYsGlyAspIleValIleMetPheGlnGlyAla 484
QY 1450 ACGGAGAGAGAGAGTACTGAGCGCGGAGATGAAATATCGGGCTTACTTGTCTGCA 1509
DB 485 ThrGlnGlnGlnYsLeuLeuSerArgYsMetMetYsTrpAlaThrPheAlaArg 504
QY 1510 ACCGGAAATCCTAATGAGAAACGACTGTCTGTGGCCAGCTTATATCTGACTGAGCAG 1569
DB 505 ThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaYrAsnLeuThrGlnGln 524
QY 1570 TACTCCAGTGTGACTTGAACATGAGCTCTGACAGAGATCTCAAGAACCGCGGTGGAT 1629
DB 525 TyrLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuYsGlnProArgAsp 544
QY 1630 TTTTGACACAGACCATCCCC 1650
DB 545 ValTrpValThrGlyTrpPro 551

```

RESULT 9
 ADR19663
 ID ADR19663 standard; protein; 618 AA.
 XX
 AC ADR19663;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Human drug metabolising enzyme (DME) -2 protein sequence.
 XX
 KW drug metabolising enzyme; DME; cytostatic; immunosuppressive;
 KW antiinflammatory; endocrine; ophthalmological; gastrointestinal;
 KW hepatocarcinoma; cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; endocrine disorder; eye disorder;
 KW gastrointestinal disorder; liver disorder; metabolic disorder; enzyme;
 human.
 XX
 OS Homo sapiens.
 XX
 PN MO200226988-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US030662.
 XX
 PR 29-SEP-2000; 2000US-0236947P.
 PR 06-OCT-2000; 2000US-0238664P.
 PR 20-OCT-2000; 2000US-0242323P.
 PR 09-NOV-2000; 2000US-0247581P.
 PR 16-NOV-2000; 2000US-0249519P.
 PR 22-NOV-2000; 2000US-0252834P.
 PR 30-NOV-2000; 2000US-0250567P.
 XX

DB 569 MetSerLeuGlyGlnArgLeuLysGluProArgArgAspValTrpValThrGlyTrpPro 568
RESURF 10
ADB64065
ID ADB64065 standard; protein; 469 AA.
XX ADB64065;
AC
XX
DT 04-DEC-2003 (first entry)
DE Human protein encoded by clone BRAMH20021910.
XX
XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW cell regeneration; membrane protein; signal transduction-related protein;
KW transcription-related protein; osteoporosis; neurological disease;
KW cancer; tumour.
XX
OS Homo sapiens.
XX
XX EPI308459-A2.
XX
PD 07-MAY-2003.
XX
PE 28-MAR-2002; 2002EP-00007401.
XX
PR 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002US-00350978.
XX
XX
PA (HELI-) HELIX RES INST.
XX (REMS-) RES ASSOC BIOTECHNOLOGY.
XX
P1 Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
P1 Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
P1 Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
XX MPI: 2003-450961/43.
DR N-PSDB; ADB62095.
XX
PT New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
PS Claim 1; Page: 222pp; English.
XX
XX The invention discloses a polynucleotide comprising a sequence selected
XX from 1970 fully defined nucleotide sequences which encode novel
XX polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
XX or its partial peptide, an antibody binding to the polypeptide or peptide
XX of the polynucleotide, immunologically assaying the polypeptide or
XX peptide of the polynucleotide by contacting the polypeptide or peptide
XX with the antibody of the encoded protein, and observing the binding
XX between the two, a transformant carrying the polynucleotide in an
XX expressible manner and an antisense polynucleotide. The oligonucleotide
XX is useful as a primer for synthesizing the polynucleotide, or as a probe
XX for detecting the polynucleotide. The polynucleotides and encoded
XX proteins are useful as pharmaceutical agents and many disease-related
XX genes may be included in them, for developing a diagnostic marker or
XX medicines for regulation of their expression and activity, or as targets
XX of gene therapy. The genes are involved in tissue and/or cell
XX regeneration. Membrane proteins, signal transduction-related proteins,
XX transcription-related proteins, disease-related proteins and genes
XX encoding them can be used as indicators for diseases (e.g. osteoporosis,
XX neurological diseases, cancer, tumours. The cDNA may be used to regulate
XX the activity or expression of the encoded protein to treat diseases. The
XX sequence presented is a protein of the invention. Note: Some of the
XX sequence data for this patent is not represented in the printed
XX specification, but is based on sequence information supplied by the
XX European Patent Office.
XX
XX Sequence 469 AA:
XX
XX Alignment Scores: 5.42e-227 Length: 469
XX
XX Pred. No.: 5.42e-227 Length: 469

Score: 2454.00
Percent Similarity: 100.00%
Best Local Similarity: 99.57%
Query Match: 77.41%
DB: 7
Gaps: 0
US-10-023-515-3 (1-1746) x ADB64065 (1-469)
QY 337 ATGCTCAAGGTGATACCCGAAATTCGAGAGTGTCAAGAGCTGCTTACCTGAACATC 396
DB 1 MetLeuLysValHisIleTyrProLysPheGlyValIleSerGluAspCysIleuTyrLeuAsnIle 20
QY 397 TATGCGCTGCGCCAGCGCGATACAGGCTCCCAAGCTCCCGCTTGTTGGTGTGTTCCAGGA 456
DB 21 TyrAlaProAlaHisAlaAspThrGlySerTyrLeuProValLeuValTrpPheProGly 40
QY 457 GGTGCTTCAAGACTGCTGAGCTCCATCTTGTATGGGTCCGCGCTGAGCTTATAG 516
DB 41 GlyAlaPheLysThrGlySerAlaSerIlePheAspGlySerAlaLeuAlaIleTyrGlu 60
QY 517 GACGTGCTGTTGTGTGTCAGTACCGGCTAGAAATATTGGATTCTTACCACATGG 576
DB 61 AspValLeuValValValValGlnTyrArgLeuGlyIlePheGlyPhePheTrpThrTrp 80
QY 577 GATCAGCATGCTCCGGGGAACGCGGCTTCAAGGACAGGCTGCTGCTGCTGCTGCTC 636
DB 81 AspGlnHisAlaProGlyAsnTrpAlaPheLysAspGlnValAlaAlaLeuSerTrpVal 100
QY 637 CAGAAAGACATCGATTCTTCGTTGGGAGCCCGAGCTGTGACCATCTTGGCGAGTCC 696
DB 101 GlnLysAsnIleGluPhePheGlyGlyAspProSerSerAlaThrIlePheGlyGluSer 120
QY 697 GCGGAGACCATTAAGTGTCTTACTGTTATCTGCTTCCATGCGCAAGGCTTATTCAC 756
DB 121 AlaGlyAlaIleSerValSerIleLeuIleLeuSerProMetAlaLysGlyLeuPheHis 140
QY 757 AAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTACCTGAGGCGCATGATTAGAG 816
DB 141 LysAlaIleLeuIleGluSerGlyValAlaIleIleProTyrLeuGlnAlaHisAspTyrGlu 160
QY 817 AAGATGAGGACCTGCGAGTGTGTCACATTTCTGTGTGTAACATGCTGAGACTGAG 876
DB 161 LysSerGluAspLeuGlnValValAlaHisPheCysGlyAsnAsnAlaSerAspSerGlu 180
QY 877 GCCCTGTGAGGTGCTGAGGAGCAAAACCTCCAAAGAGCTGTGACCTGACCGAGAA 936
DB 181 AlaLeuLeuArgCysLeuLysGlnTrpLysProSerLysGluLeuLeuThrLeuSerGlnLys 200
QY 937 ACAAGTCTTTCAGTGAAGTGTGATGGTCTTCTTCTTATGAGCCTTAGATCTTA 996
DB 201 ThrLysSerPheThrArgValValAlaAspGlyAlaPhePheProAsnGluProLeuAspLeu 220
QY 997 TTGTCTCAAAAGCATTTAAAGCAATTCTTCATGATGTGAGTCAATTAACGAGTGT 1056
DB 221 LeuSerGlnLysAlaPheLysAlaIleProSerIleIleGlyValAsnSerHisGluCys 240
QY 1057 GGTTCCTGCTGCTTGAAGGAGGCTCGAGATCTGAGTGGCTCCAAAGTCCCTT 1116
DB 241 GlyPheLeuLeuProMetLysGluAlaProGlnIleLeuSerGlySerAsnLysSerLeu 260
QY 1117 GCCCTCATCTGATACAAACATCTGACATCCGCTCAGTATTGTGACCTTGAGCT 1176
DB 261 AlaLeuHisIleuIleGlnAsnIleLeuHisIleProProGlnTyrLeuHisIleuValAla 280
QY 1177 AATGAATATTCTTCATGACAAAGCATCTCCGATGAAATCCGAGAGTCTTGAAGTTG 1236
DB 281 AsnGlnTyrPheHisAspLysHisSerLeuThrGlnIleArgAspSerLeuLeuAspLeu 300
QY 1237 CTTGAGATGTGTGTTCTTGTGTGCTGCTGACATGACAGCTCGATATCAGAGAGTGT 1296
DB 301 LeuGlyAspValAlaPhePheValAlaProAlaLeuIleThrAlaArgTyrHisArgAspAla 320
QY 1297 GGTGACCTGTCTACTTCTTATGAGTTTGGGACCGGCTCAGTGTCTTGAAGACGAAAG 1356

|||||
Db 321 GlyAlaProValItyrPheTyrGluPheArgHisA-GProGlnCysPheGluPheThrIys 340
QY 1357 CCGGCTTTTTCGAAGCCGACCGCTGATGAAGTCCGCTTGTGTTGGTGTGCTTC 1416
Db 341 ProAlaPheValIyValAlaPheHisAlaPheGluValArgPheValPheGlyValAlaPhe 360
QY 1417 CTGAAGGGGAGCATGTTATGTTCTGAAGAGGACGAGAGAGAGAGATTACTGAGCCGG 1476
Db 361 LeuIySGIyAlaPheIleValMetPheGluGluIyAlaThrGluGluIyIleuLeuSerArg 380
QY 1477 AAGATGATGAATACTGAGGCTTACCTTTGCTCGAACCGGAAATCTTAATGGAGACCTTG 1536
Db 381 LysMetMetIyIyTyrTrpAlaThrPheAlaArgThrGlyAsnProAsnGlyAsnAspIeu 400
QY 1537 TCTCTGTGGCCAGCTTATATATCTGACGAGTACCTTCAGCTGAGCTTGAACATGAGC 1596
Db 401 SerLeuTrpProAlaTyrAsnLeuThrGluGlnIyLeuAspIeuAsnMetSer 420
QY 1597 CTCGAGCAGAGACTCAAGAACCGGAGGATTTTGGACGAGACCATCCCTGATC 1656
Db 421 LeuGlyGlnAlaGlyLeuGluProArgValGluPheTrpMetThrIleProIeuIle 440
QY 1657 CTGTCTGCTCGACATCTCCACAGTCTCTTCTTCTTAACTTCTCTCTCTCTC 1716
Db 441 LeuSerAlaSerAspMetLeuHisSerProIeuSerSerLeuThrPheLeuSerIeuIeu 460
QY 1717 CAGCTTTCTTTTCTTTTCTTGTGCTCT 1743
Db 461 GlnProPhePhePheCysAlaPro 469

RESULT 11

ADFS0145
ID ADFS0145 standard; protein; 542 AA.

AC ADFS0145;

DT 12-FEB-2004 (first entry)

DE Cat cauxin protein SEQ ID NO:2.

KW cat; cauxin; cat kidney disease marker; kidney disease.

OS Felis catus.

PN JP2003250575-A.

PD 09-SEP-2003.

PE 04-MAR-2002; 2002JP-00057908.

PR 04-MAR-2002; 2002JP-00057908.

PA (TOHO-) TOHOKU TECHNOARCH KK.

DR MPI: 2004-002277/01.

DR N-PSDB; ADFS0144.

PT Novel cauxin protein or its salt, useful as a cat kidney disease marker,

PS Claim 3; SEQ ID NO 2; 33pp; Japanese.

XX The present sequence represents a cat cauxin protein (I) or its salt,
CC which is cat kidney disease marker. Also described: (1) a partial peptide
CC (II) of (I); (2) a nucleic acid (III) encoding (I) or (II); (3) a vector
CC (IV) comprising (III); (4) a transformed host (V) comprising (III) or
CC (IV); (5) producing (I) or (II) by culturing (V); (6) an antibody (VI)
CC which couples specifically with (I) or (II); (7) diagnosing cat kidney
CC disease which involves measuring (I) quantitatively, and where reduction
CC of amount of (I) indicates presence of the disease; (8) a cat kidney
CC disease diagnostic agent comprising (I) labelling agent, a reagent which
CC measures the biological activity of urinary (I) or (VI); and (9) a cauxin

CC detection kit which measures cauxin in a test sample. (I) is useful as a
CC cat kidney disease marker and (VI) is useful for diagnosing cat kidney
CC disease. (I) enables detection of cat kidney disease simply and
CC correctly (I) provides an early marker for the disease, and replaces
CC complicated diagnostic methods such as X-ray imaging, ultrasonic imaging,
CC blood testing.

XX SQ Sequence 542 AA;

Alignment Scores:

Pred. No.:	1e-181	Length:	542
Score:	1985.00	Matches:	373
Percent Similarity:	83.27%	Conservative:	60
Best Local Similarity:	71.73%	Mismatches:	85
Query Match:	62.62%	Indels:	2
DB:	8	Gaps:	2

US-10-023-515-3 (1-1746) x ADFS0145 (1-542)

QY 91 GGGCTTGTGCTGAAGGGCCACAGAGAACAGAGCTGGAGTTCAGGGCAAGCA 150
Db 22 GlyProAlaAlaPheAlaProValArgSerThrArgLeuGlyTyrValArgGlySerGln 41
QY 151 GTCACTGTGTGGAAAGCCGTGTGCTGTGAACGTGTCTTGTGAGTCCCTTGTGCT 210
Db 42 ThrThrValIeuGlySerThrValProValIeuMetPheLeuGlyIleProTyrAlaAla 61
QY 211 CCCCCTGGGATCCCTCGATTAACGAACCCGAGCTCGACATCGCCCTGGATACCTG 270
Db 62 ProProIeuGlyProIeuAlaArgPheIySerGlnProIyAlaIeuProGlyIyAsnAspPhe 81
QY 271 CGAGAGCCACCTCTTACCTTAATTTGTGCTCCGAGACTGAGAGTGGCTGCTTAAGT 330
Db 82 ArgAsnAlaTrpSerTyrProIyIeuCysPheGlnIlePheLeuGlyTyrIleValSerTyr 101
QY 331 CAACATGCTCAAGGTGCTATTAACCCGAATTCGAGAGTGCAGAGACTGCTTACCTG 390
Db 102 GlnHisValIeuIyValArgTyrProIyIeuGlnIyAsnSerGlnIyAspCysLeuTyrIeu 121
QY 391 AACATCTATGCGCTGCCCGACCGGATACAGGCTTCGAAGCTCCCGCTTGTGTGCTTC 450
Db 122 AsnIleTyrAlaProAlaHisAlaAspAsnGlySerAsnIleuProValMetValTrpPhe 141
QY 451 CCAGAGGTGCTTCAAGACTGCGTCCAGCTCCATCTTGAATGGGTCCGCTGGCTGCC 510
Db 142 ProGlyIyAlaPheIyMetGlySerAlaSerPheAspGlySerAlaIeuAlaIa 161
QY 511 TATGAGACGTGTGTTGTGTGCTCCAGTACCGGCTAGGAAATTTGTTCTTCAAC 570
Db 162 TyrGluAspValIeuIleValThrThrIleTyrTrpIeuGlyIlePheGlyPhePheAsp 181
QY 571 ACATGGATCAAGATCTCCGGGAAGTGGCTTCAAGAACAGAGTGGCTGTCTGCC 630
Db 182 ThrGlyAspGlyHisAlaArgGlyAsnTrpAlaIeuIeuAspGlnValAlaIeuThr 201
QY 631 TGGTCCAGAGAACATCGAGTTCTGTGTTGGGAGCCCACTGTGTAACATCTTGGC 690
Db 202 TrpValArgAspAsnIleGlnPhePheGlyIyAspProArgSerValThrIlePheGly 221
QY 691 GAGTCCGGGAGGACATTAAGTTTCTAGTCTTACTAGTCTCCCATGGCAAGGCTTA 750
Db 222 GluSerAlaGlyAlaIleSerValSerSerIeuIleIeuSerProIleAlaIeuIyIeu 241
QY 751 TTCACAAAGCCATCATGAGAGTGGGTGGCCATCATCTTCACTG--GAGGCCAT 807
Db 242 PheHisIyValAlaIleMetGluSerGlyValAlaIleIeuProIeuIeuMetArgProPro 261
QY 808 GATATATGAGAGATGAGAGTCAAGTGTGTTGCTTCACTTCTGTGTAAGTAAAGTGGCA 867
Db 262 GlyAspGlnIyGlyIyAspIeuGlnIyAlaIeuIyAlaArgIleCysGlyIyCysHisIleAsp 281
QY 868 GACTCTGAGGCTCGCTGAGTGGCTGAGGACAAACCTCCAGAGACTGTGACCTG 927

631 TGGTCCAGAAAGATGAGTTCTTGGTGGGAGCCCAAGCTGTGACCATCTTTGGC 690
DB 202 TTPVALAAGAAPHANILEGLUPHEHEGLYLAAPROAARGSERVALTHIRLEHEGLY 221
QY 691 GAGTCGGGGAGGACATAGTGTCTTACTTATAGTCTCCATGCGCAAGGCTTA 750
DB 222 GUSERIALAGLYALALESERVALSERLEULEUSERPROLLEALLANGLYUEN 241
QY 751 TTCCCAAAAGCCATATAGAGATGGGGTGGGCATATCCCTTACTCTG---GAGGCCAT 807
DB 242 PHEHSLYVALALEMGLUSERGLYVALAALILEUENPROLEUENUEHARGPRO 261
QY 808 GATATAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 867
DB 262 GLYAPGLUARGLYLYAPAPHEUENGLINVALLEUALLAAGLYCYEHISALASER 281
QY 868 GACTCTGAGGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 927
DB 282 AEPSEVALAALALEUENGLINCYSEUARGALYSPROSERGLUGLEUENUEHARG 301
QY 928 AGCCGAAAGAAAGCTTTTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 987
DB 302 SERLYSLYLEUTHIRPHESEVALLEPROVALLEAPAPAPHEPHEPHEPHEPHEPHE 321
QY 988 CTAGATCTATGCTCTCAGAAAGCATTTAAAGCAATTCCTTCATGAGGAGGAGGAGGAG 1047
DB 322 VALAALALEUENGLINGLYALAPHEANSEVALPROSERILEGLYALANSEN 341
QY 1048 CACGAGTGTGCTTCTGCTGCTTGAAGAGGCTCTGAGATCTGAGGAGGAGGAGGAG 1107
DB 342 HISGLUCYBALAPHELEUEN---SERTHIRGLIPHESEGLULILEUGLYGLYSEHEN 360
QY 1108 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1167
DB 361 ARGSELEUVALALEUENGLYVALHISTHRPHLEUENGLYLEPROTHGLINYLEUENHIS 380
QY 1168 CTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1227
DB 381 LEUVALAALAPHEHISRYRPHERYRANGLYSHISSEPROVALGLULILEUENUEHARG 400
QY 1228 CTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1287
DB 401 LEUAPHELEUENGLYAPHEVALLEUENHISVALPROGLYVALVALTHIRLARGYHIS 420
QY 1288 AGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1347
DB 421 ARGAPALAGLYALAPROVALYRPHERYRGLUPHEGLINHISLEPROGLINCYSEUEN 440
QY 1348 GACGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1407
DB 441 AAPHTRARGPROALAPHEVALLYSALAPHEHISSEARPGULILEARGPHEVALHEGLY 460
QY 1408 GGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1467
DB 461 GLYALAPHELEUENGLYAPHELEVALMETPHEGLUGLYALATHGLUGLUGLYUEN 480
QY 1468 CTGAGCGGAGAGATGATAATATCTGGCTACCTTGTCTGCAACCGGAGATCTTAATGG 1527
DB 481 LEUSERARGLYMEETECARGYRTRPALAENPHEALAAARGTHIRGLYAPPROANGLY 500
QY 1528 AACGACCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1587
DB 501 GLUGLYVALPROLETRIPROALATYRTHIRGLINSEARGUGLINTYRLEUENUEHARG 520
QY 1588 AACATGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1647
DB 521 SERVALSERVALGLYGLINLYSEUENGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 540

AC ABP61006;
XX 10-SEP-2002 (first entry)
DT
XX
DE Novel human protein. SEQ ID 93.
XX
KW Human; cytosolic; vulnary; antiarteriosclerotic; antiparkinsonian;
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antilucer; vitruide; antihypoid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyperimmunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty.
XX
OS Homo sapiens.
XX
PN WO020250105-A1.
XX 27-JUN-2002.
PD
XX
PF 17-DEC-2001; 2001WO-US049232.
XX
PR 19-DEC-2000; 2000US-0256710P.
XX 20-DEC-2000; 2000US-0257048P.
PR 09-JAN-2001; 2001US-0260482P.
PR 30-JAN-2001; 2001US-0264922P.
PR 06-FEB-2001; 2001US-0266797P.
PR 19-MAR-2001; 2001US-0276988P.
PR 04-APR-2001; 2001US-0281535P.
PR 08-MAY-2001; 2001US-0289622P.
XX
XX (SMK) SMITHKLINE BEECHAM CORP.
XX (SMK) SMITHKLINE BEECHAM PLC.
XX (GLAX) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kahnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
DR WPI; 2002-508784/54.
XX N-PSDB; AB086171.
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune disorder.
XX
PS Claim 1(a); Page 315; 335pp; English.
XX
CC The invention relates to an isolated polypeptide with signal sequences
CC which allow it to be secreted extracellularly or membrane associated. The
CC activity of polypeptides of the invention may be described as,
CC cytosolic, vulnary, antiarteriosclerotic, antiparkinsonian, neurotropic,
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
CC cardiant, antilucer, vitruide, antihypoid, cerebroprotective, anorectic,
CC and metabolic. Polypeptides and polynucleotides of the invention are
CC useful in the treatment, or as a vaccine in the prevention of, cancer,
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
CC inflammation, neoplastic diseases, nervous system related disorders and
CC cardiovascular disorders, pancreatitis, respiratory disorder.
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,
CC developmental abnormality, gastrointestinal ulceration, neuropathy,
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid
CC disorders e.g. hypothyroidism, brain damage, colitis, cone photo-
CC transduction deficiency, neurological diseases, stroke, angiogenesis,

CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
 CC trachea, thymus, lymph node and muscular system, obesity, anorexia,
 CC growth abnormalities and alleviation of precocious puberty. The
 CC sequences given in records ABP60965-ABP61019 represent novel human
 CC proteins of the invention

XX Sequence 356 AA:

Alignment Scores:

Pred. No.:	4,7e-162	Length:	356
Score:	1780.00	Matches:	340
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	56.15%	Indels:	0
DB:	5	Gaps:	0

US-10-023-515-3 (1-1746) x ABP61006 (1-356)

```

QY 724 ATACTGCTCCGAGCCCAAGGCTTATTCACAAAGCCATGAGAGTGGGCTGCC 783
D 17 ILeuSerProMetAlaIysGlyLeuPheHISlyValaIleMetGluSerGlyValaIa 36
QY 784 ATCATCCCTTACCTGAGGCCCCATGATTATGAGAAAGTGAAGACCTGAGGTGTTGCA 843
D 37 ILeIleProTyLLeuGluValaHISaAspTyrgIuYsSerGluAspLeuGluValaIa 56
QY 844 CATTTCTGTGTAACATGCGCTCAGACTCTGAGGCGCTGAGGCTGAGGCTGAGCAAA 903
D 57 HisPheCyseGlyAsnAsnHisSerAspSerGluValaLeuLeuAspGlyLeuGlyHis 76
QY 904 CCCTTCAGAGAGCTGCTGACCTCAGCCAGCAAAACAAAGTCTTTGACTGAGTGGTTGAT 963
D 77 ProSerIlysgIuLeuThrLeuSerGlnYsThrIlySerPheThrArgValaIAsp 96
QY 964 GATGCTTCTTTCCATAGAGCGCTTAGATCTATGCTCAGAAAGCATTTAAAGCAATT 1023
D 97 GlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnYValaIaPheValaIle 116
QY 1024 CTTTCATCATCGAGTCAATTAACACAGATGTGGCTTCTGCTGCTATGAAGAGGCT 1083
D 117 ProSerIleIleGlyValaAsnAsnHisGluGlyGlyPheLeuLeuProMetIlyGluAa 136
QY 1084 CCTGAGATCTCTCATGAGGCTCCAAACAAGTCCCTTGGCTTCATCTGATTCAAACAATCCTG 1143
D 137 ProGluIleLeuSerGlySerAsnIlySerIleuValaLeuIleGlnAsnIleLeu 156
QY 1144 CACATCCCGGCTCAGTATTTGACCTTGGCTATGAATACTTCCTCAATGACAGACATCC 1203
D 157 HisIleProProGlnIlyLeuHISleuValaIaAsnGlnIlyrPheHISAspIlyHisSer 176
QY 1204 CTGACTGAATCCGAGACAGTCTTCTGACTTGTGGAGATGTGTTCTTTGTGGTCCCT 1263
D 177 LeuThrGluIleAspSerSerIleuAspLeuLeuGlyAspValaPhePheValaIPro 196
QY 1264 GCACGTATCAGGTCCATATCAAGAGATGCTGTGACCTGTCTACTTTATGAGATT 1323
D 197 AlaLeuIleThrAlaArgIyHISaAspIlyAspIlyAlaProValaIlyrPheTyGluPhe 216
QY 1324 CGGACCGGCTCAGTCTTTGAGACAGACGCGGCTTTTGTCAAGCGACAGCAGCT 1383
D 217 ArgHISaArgProGlnIlyrPheGluAspThrIlyrProAlaPheValaIlyAspHisAla 236
QY 1384 GATGAAGTCCGCTTGTGTGTGCTGTGGTCTTCTGTAAGGGAGCATTTGTTATGTTGAA 1443
D 237 AspIlyValaIlyrPheValaIlePheGlyGlyAlaPheLeuIlysgIyAspIlyValaIleMetPheGlu 256
QY 1444 GAGGCCAGGAGAGAGAGATTACTAGCCGGAAGATGATGAATCTGGGCTACTTT 1503
D 257 GlyAlaIleThrIlyGluGluIlyLeuLeuSerArgIlySerMetCtyIyTrpAlaIlePhe 276
QY 1504 GCTTCGAACCGGAGATCTTAATGGAGACAGCTGTCTGTGGCGACTTAATCTGACT 1563
D 277 AlaArgThrIlyAsnProAsnGlyAsnAspLeuSerIleuTrpProAlaIlyrAsnLeuThr 296

```

```

QY 1564 GAGCAGTACCTTCACAGCTGACTTGAACATGAGCCTCGAGACAGAGACTCAAGAACCCGG 1623
D 297 GluGlnTyLeuGlnIleuAspLeuAsnMetSerLeuGlnIlyAspGlyGluProArg 316
QY 1624 GTGATTTTGGACAGACAGCATCCCTGATCCGTGCTCGCATCTCCACAGT 1683
D 317 ValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHisSer 336
QY 1684 CTTCTTTCTTCTTAACCTTCTCTCTCTCTCCAGGCTTTCTTTTGTGTGCTCT 1743
D 337 ProLeuSerSerIleuThrPheLeuSerIleuGlnIProPhePhePheCysAlaIPro 356

```

RESUR 14
 AAE20911
 ID AAE20911 standard; protein; 561 AA.
 AC AAE20911;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Rat carboxylesterase-like enzyme protein.
 XX
 KW Rat; carboxylesterase-like enzyme; organophosphorus intoxication;
 KW osteopathic; gene therapy; osteoporosis; antitense therapy; cytostratic;
 KW detoxifying agent; Paget's disease; bone implant degradation; cancer;
 KW dental implant; enzyme; gene expression.
 XX
 OS Rattus sp.
 XX
 PN M0200206454-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 10-JUL-2001; 2001WO-BP007919.
 XX
 PR 17-JUL-2000; 2000US-0218564P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Xiao Y;
 DR WPI; 2002-195808/25.
 XX
 PT Novel human carboxylesterase-like enzyme polypeptide, regulators of which
 PT are useful for preventing and treating organophosphorus intoxication,
 PT cancer and osteoporosis.
 XX
 PS Disclosure; Fig 8; 92pp; English.
 XX
 CC The invention relates to a purified human carboxylesterase-like enzyme
 CC polypeptide. Carboxylesterase-like enzyme and its DNA are useful for
 CC screening for agents which decrease or modulate the activity of
 CC carboxylesterase-like enzyme polypeptide. Carboxylesterase-like enzyme is
 CC useful for treating a carboxylesterase-like enzyme dysfunction related
 CC disease, such as organophosphorus intoxication, cancer and osteoporosis.
 CC Compounds that increase the ability of human carboxylesterase-like enzyme
 CC to bind to organophosphorus compounds are useful as detoxifying agents.
 CC Carboxylesterase-like enzyme agonists and antagonists are useful for
 CC treating osteoporosis. Paget's disease and degradation of bone implants,
 CC particularly dental implants. Carboxylesterase-like enzyme is useful in
 CC diagnostic assays for detecting diseases and abnormalities or
 CC susceptibility to diseases or abnormalities related to the presence of
 CC mutations in the nucleic acid sequences which encode the enzyme. The
 CC coding sequence of carboxylesterase-like enzyme polynucleotide is useful
 CC in gene therapy and for generating antisense oligonucleotides or
 CC ribozymes which specifically bind to mRNA transcribed from
 CC carboxylesterase-like enzyme DNA. These antisense oligonucleotides are
 CC useful for modulating carboxylesterase-like enzyme gene expression. The
 CC present sequence is rat carboxylesterase-like enzyme protein
 CC
 XX
 SQ Sequence 561 AA;

Alignment Scores:

Pred. No.: 4,886-112 Length: 561
 Score: 1262.50 Matches: 275
 Percent Similarity: 62.41% Conservative: 77
 Best Local Similarity: 48.76% Mismatches: 179
 Query Match: 39.83% Indels: 33
 Gaps: 13

US-10-023-515-3 (1-1746) x AAB58981 (1-561)

```

OY 1 ATGCCACAGGGA---CTTACTTCATCTGCTTCACAAATGCTTTTCTGATTTCTCCAG 57
Db 1 MetProArganGlnLeuHisSerTrpLeuAsnAlaValLeuPheGlyLeuLeuLeu--- 19
OY 58 CCCGTTGGGACACAGACAGTGGGCAAAAACCTGGCCCTTCCTCGAAGGGCCACAGAG 117
Db 20 ---LeuLeuL1LeuLsValGln---GlyGlnSerProGluSer---SerProL1LeuArg 36
OY 118 AACACACAGCTGGATGGATTCAGGGCAAGCAATCACTGTGCTGGAGCCCTGTGCT 177
Db 37 ThrThrHisThrGlnGlnValGlnGlnIlyLeuAspHisValLysAspThrIlyAlaGly 56
OY 178 GTGAACGTGTTCTTGGAGTCCCTTTGCTGCTCCCGCTGGAGATCCCTGGATTACG 237
Db 57 ValHisThrPheLeuGlyIlyProPheAlaLysProValGlyProLeuArgPheAla 76
OY 238 AACCCGACGCTGCATCCGCTGGGATTAATTCGACAGAACCCACTCTCACTAATTTG 297
Db 77 ProProGlnProProGlnProGlnProTrpSerGlyValArgAspAlaThrSerGlnProAlaMet 96
OY 298 TGCCTCCAGACTAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 345
Db 97 CysLeuGlnAsnLeuAspL1LeuAspGlnValGlyLeuLeuAspMetLysMetLys--- 115
OY 346 GTGCATTAACCCGAAATTCGAGTGCAGAGAATGCTCTTACCTGAACATGTATGCGCT 405
Db 116 -----LeuSerSerL1SerMetSerGlnAspCysLeuGlyLeuAsnValIlyAlaPro 133
OY 406 GCCCAGCCGATACAGGCTCCAGGCTCCGCTGTTGGTGGTGGTCCCGAGAGTCCCTTC 465
Db 134 AlaHisAlaArgGlnGlySerAsnLeuProValMetValTrpL1LeuGlyAlaLeu 153
OY 466 AAGATGCTAGGCTTCATCTTGTATGGATGCGCCGCTGGCTGCTGCTGCTGCTGCTG 525
Db 154 ValValGlyMetAlaSerMetLysTrpAspLysLeuLeuThrValAsnGlyAspLeuVal 173
OY 526 GTTGTGTCGTCCAGTACCGGCTAGAGAAATTTGGTTTCTTCACCAATGGAGTACGAT 585
Db 174 ValValThrL1LeuGlnIlyArgLeuGlyValLeuGlyPhePheSerThrGlyAspGlnHis 193
OY 586 GCTCCGGGAAACTGGGCTTCAGAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
Db 194 AlaArgGlyAsnTrpGlyIlyLeuAspGlnValAlaAlaLeuArgTrpValGlnGlnAsn 213
OY 646 ATCGAGTCTTCGGTGGGAGACCCGCTGTCGACATCTTTGGAGAGTCCGCGGAGAGCC 705
Db 214 L1LeuAsnL1PheGlyGlyAsnProAsnArgValTrpL1PheGlyGlnSerAlaGly 233
OY 706 ATAAAGTCTTCTAGCTTATATCTCTCCATGCGCAAGGCTTATTCACAAAGCCATC 765
Db 234 ThrSerValSerSerHisValL1SerProMetSerGlnGlyLeuPheHisGlyAlaLys 253
OY 766 ATGAGAGTGGGGTGGCCATCATCTCCCTTACCTGGAGGCCCATGATTATGAGAAAGTGA 825
Db 254 MetGlnSerGlyValAlaLeuLeuProAspLeuL1Ser-----GlnThrSerGln 270
OY 826 GACCGCAGG---GTGGTTCACATTTCTGTGTAAACATGGCTGACAGCTGAGGCCCTG 882
Db 271 ThrValSerThrThrValAlaLysLeuSerGlyCysGlyAlaMetAspSerGlnLysLeu 290
OY 883 CTGAGAGTCTGAGAGCAAAACCTTCACAGAGTCTGAGCCCTGACGACAGAAACAAAG 942
Db 291 ValArgCysLeuArgAlaLysSerGlyAlaGlnLysLeuValL1AsnLysValPheLys 310
  
```

```

OY 943 TCTTCACTGAGTGGATGATGCTTCTTCTTCTTAATGACCTCTAGATCTATTTCT 1002
Db 311 MetL1ProAlaValAlaAspGlyGlnPheLeuProArgHisProLysGlyLeuLeuAla 330
OY 1003 CAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACACAGAGTGGCTTC 1062
Db 331 SerGlnAspPheHisProValProSerL1LeuGlyValAsnThrAspGlnIlyCysCys 350
OY 1063 CTGCTGCTATG-----AAGAGGCTCCCGAGTCTCTGCT 1098
Db 351 ThrL1LeuProMetValMetGlyThrAlaGlnL1LeuLysGln-----LeuSer 366
OY 1099 GAGTCCACAAAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158
Db 367 ArgGlyAsnLeuGlnAlaValLeuLysAspHisAlaAlaGlnMetMetLeuProGln 386
OY 1159 TATTTGACCTTTGGCTTAATGAAATCTTCATGACAAAGCACTCCCTGATGAAATCCGA 1218
Db 387 CysGlyAspLeuLeuMetGlnIlyTrpMetGlyAsnThrAspAspSerGlnThrLeuGln 406
OY 1219 GACAGTCTTTCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1278
Db 407 L1LeuGlnIlyThrGlnMetMetGlyAspPheLeuPheValL1ProAlaLeuGlnValAla 426
OY 1279 CGATATGACAGAGATGCTGTCACCTGCTTACTTCTTGAATTCGAGCTGAGTGGCTGAG 1338
Db 427 HisPheGlnArg---SerHisAlaProValIlyPheIlyGlnPheGlnHisAlaProSer 445
OY 1339 TGCCTTGAAGACACAGAACCCGCTTTTGTCAAAGCCGACCACTGATGAAGTCCGCTTT 1398
Db 446 TyrPheLysAsnValArgProHisValLysAlaAspHisAlaAspGlyValL1ProPhe 465
OY 1399 GTTGTGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1458
Db 466 ValPhe---GlySerPhePheSerGlyMetLysLeuAspPhe-----ThrGlnGln 481
OY 1459 GAGAGTTCATGAGCCGAGATGATGAATACTGGGCTTACTTGTCTGCAACCGGAGAT 1518
Db 482 GlnTrpLeuLeuSerAspArgMetMetLysTrpTrpAlaAsnPheAlaArgGlnGlnAsn 501
OY 1519 CCTAATGGAACGACCTGCTCTGTGGCCAGCTTATTAATCTGACTGACAGTACCTCGAG 1578
Db 502 ProAsnSerGlnGlyLeuProIlyTrpProAlaLeuAspHisAspGlnGlnIlyLeuGln 521
OY 1579 CTGGACTTGAACATGAGGCTCGACAGACATCAAGAAACCGGGTGGATTTTGGACC 1638
Db 522 LeuAspThrHisProAlaValAspArgAlaLeuLysAlaArgArgLeuGlnPheTrpThr 541
OY 1639 AGCACCATCCCC 1650
Db 542 LysThrLeuPro 545
  
```

RESULT 15
 AAB58981 standard; protein; 549 AA.
 AAB58981;
 27-MAR-2001 (first entry)
 Breast and ovarian cancer associated antigen protein sequence SEQ ID 689.
 Human; breast cancer; cytotoxic; immunosuppressive;
 KM nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KM antidiabetic; antiinflammatory; anticancer; vulnery; anticonvulsant;
 KM antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KM Addison's disease; allergy; autoimmune haemolytic anemia;
 KM autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KM multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KM cardiovascular disorder; wound healing; neurological disease.
 OS Homo sapiens.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 06:53:49 ; Search time 31.53 Seconds
(without alignment)
10656.177 Million cell updates/sec

Title: US-10-023-515-3
Perfect score: 3170
Sequence: 1 atgcacagagactactctc.....cttctcttgcctcttga 1746

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO_epool/6664091/runat_14062005_133459_9346/app_query.fasta_1.4238
-DB=FIR_79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=-1 -END=-1 -MATRIX=blomsum62 -TRANS=human0.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091.qcgn.1.1_101.qrunat_14062005_133459_9346 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1298.5	41.0	557	2	A47162
2	1231.5	38.8	532	2	A43329
3	1230	38.8	559	1	JC5408
4	1199	37.8	561	2	S47655
5	1186.5	37.4	554	1	S34607
6	1163	36.7	554	2	A39060
7	1162.5	36.7	566	2	S19307
8	1144	36.1	549	2	JX0054
9	1140	36.0	567	1	A41010
10	1139	35.9	561	2	S62788
11	1138.5	35.9	562	2	A55281
12	1136	35.8	540	2	A31584
13	1129.5	35.6	561	2	S71597
14	1123	35.4	561	2	JC2447

15	1121.5	35.4	565	2	S10367	carboxylesterase (
16	1034.5	32.6	539	2	A29923	carboxylesterase (
17	777	24.5	596	1	ACRYE	acetylcholinestera
18	764.5	24.1	614	2	A39256	acetylcholinestera
19	763	24.1	614	2	JH0314	acetylcholinestera
20	756	23.8	599	1	A38868	acetylcholinestera
21	754	23.8	604	2	JH0811	acetylcholinestera
22	752	23.7	602	1	ACHU	cholinesterase (EC
23	740	23.3	603	2	S70849	cholinesterase (EC
24	732	23.1	581	2	C39768	cholinesterase (EC
25	726.5	22.9	584	2	S48724	acetylcholinestera
26	710.5	22.4	583	2	S10712	acetylcholinestera
27	692.5	21.8	620	2	A5413	acetylcholinestera
28	669.5	21.1	745	2	S13586	triacylglycerol li
29	658.5	20.8	489	2	B69680	para-nitrobenzyl e
30	643.5	20.3	597	2	A33668	sterol esterase (E
31	633.5	20.0	599	2	A57701	sterol esterase (E
32	627.5	19.8	664	2	UC7990	acetylcholinestera
33	618	19.5	612	2	A34967	sterol esterase (E
34	594.5	18.8	691	2	JE0150	acetylcholinestera
35	593	18.7	550	1	A34576	crystal protein pr
36	588	18.5	767	2	S47639	acetylcholinestera
37	581.5	18.3	746	2	A25363	acetylcholinestera
38	575	18.1	540	2	S51043	carboxylesterase (
39	572.5	18.1	602	2	T37254	acetylcholinestera
40	570.5	18.0	637	2	S66236	acetylcholinestera
41	569.5	18.0	554	2	T31783	hypothetical prote
42	569.5	18.0	557	2	A56690	esterase - Caenorh
43	562.5	17.7	545	2	A89046	protein B0238.1 li
44	557.5	17.6	593	1	S25062	triacylglycerol li
45	551	17.4	562	2	S27782	esterase precursor

ALIGNMENTS

RESULT 1

A47162

Cholinesterase B (EC 3.-.-.-) (precursor - mallard

C) (Date: 05-May-1995 #sequence_revision 05-May-1995 #text_change 09-Jul-2004

C) (Accession: A47162

R) (Hwang, C.S.; Kolatukudy, P.E.

J. Biol. Chem. 268, 14278-14284, 1993

A) (Title: Molecular cloning and sequencing of cholinesterase B cDNA and stimulation of expe

A) (Reference number: A47162; MUID:93300823; PMID:8314791

A) (Accession: A47162

A) (Status: preliminary

A) (Molecule type: mRNA

A) (Residues: 1-557 <HMA>

A) (Cross-references: UNIPROT:004791; GB:L05493; NID:9213100; PIDN:AAA49223.1; PID:9213101

C) (Superfamily: cholinesterase; cholinesterase homology

F) (56-545/Domain: cholinesterase homology <CHE>

Alignment Scores:

Pred. No.:	1,996-92	Length:	557
Score:	1298.50	Matches:	265
Percent Similarity:	62.77%	Conservative:	79
Best Local Similarity:	48.36%	Mismatches:	177
Query Match:	40.96%	Indels:	27
DB:	2	Gaps:	8

US-10-023-515-3 (1-1746) x A47162 (1-557)

QY	88	ACTGGGCTTGTGCTGAAGGGCCACAGAAACACCAAGCTGGATTCAGGCGAAG	147
DB	23	ThrglyglnlyrAlaaglglInprogluValThrAmyrGlySerValArglyr	42
QY	148	CAAGTCATGCTGCTGGAGACCTGTGCTGCAACGTCTCTGGAGTCCCTTGT	207
DB	43	GlnVallyrValAlaAlaIgluArgSerValAlaValPheLeuGlyLeuProPheAla	62
QY	208	GCTCCCGCGCTGGATCCCTCGATTTACGAACCCGACGCTGCATCCCTGGATTAAC	267

```

Db      63  LysProProValGlyProLeuArgPheSerGluProGlnProProGluProIleuProIleuValGly 82
      |||||:::|||||:::|||||:::|||||
Qy      268  TTGCGAAGAACCACTCTTAACCTTAATTTGGCTTCCAGACTCAGAGTGCGCTCTTA 327
      :::::|||||:::|||||
Db      83  ValArgAspAlaIleSerIleProPheMetCysLeuGlnAspLys-----ValLeu 99
      :::::|||||
Qy      328  GATCAACACATGCTC-----AAGGTGCATTACCCGGAATTCGAGAGTCA 372
      ||:::|||||
Db      100  GlyGlnTyrLeuSerAspAlaIleThrAsnArgLysGluValAlaArgLeuGlnIleSer 119
      :::::|||||
Qy      373  GAAAGCTGCTTATCTGAACATTAATGCGCTCCAGCCGATACAGCTCCAAAGCTC 432
      :::::|||||
Db      120  GlnAspCysLeuTyrLeuAsnValTyrThrProValSerThrGlnGluGlnIleuVal 139
      :::::|||||
Qy      433  CCGGCTTGTTGGTGGTCCAGAGAGTGCTTCAAGACTGCTCAGCTCCATCTTTGAT 492
      |||||:::|||||
Db      140  ProValPheValTyrIleHisGlyGlyLeuValSerGlyAlaIleSerSerIleAsp 159
      :::::|||||
Qy      493  GGGTCCGCTGCTGCTGCTATGAGAGAGTGTGTGTGTGCTCCAGTACCGGCTAGGA 552
      |||||:::|||||
Db      160  GlySerAlaLeuAlaIlePheAspAsnValValValThrIleGlnTyrArgLeuGly 179
      :::::|||||
Qy      553  ATATTGGTTTCTTACCAATGAGATGAGATGCTCCGGGGAAGTGGCTTCAAGAGC 612
      :::::|||||
Db      180  IleAlaGlyTyrPheSerThrGlyAspLysHisAlaArgGlyAsnTyrGlyTyrLeuAsp 199
      :::::|||||
Qy      613  CAGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
      |||||:::|||||
Db      200  GlnValAlaAlaLeuGlnTyrIleGlnGlnAsnIleIleHisPheAspGlyAspProIle 219
      :::::|||||
Qy      673  TCTGTGACCATCTTTTGGAGAGTCCGCGGAGCCATAGTGTGTGTGCTTATCTGCT 732
      :::::|||||
Db      220  SerValThrIlePheGlyGlySerAlaGlyValSerValSerAlaLeuValLeuSer 239
      :::::|||||
Qy      733  CCCATGGCCAAAGCTTATTCACAAAGCCATCAGAGAGTGGGCTGGCCATC---ATC 789
      |||||:::|||||
Db      240  ProLeuAlaLysGlyLeuPheHisIleValAlaIleSerGluSerGlyThrAlaValArgIle 259
      :::::|||||
Qy      790  CCTTACCTGAGAGCCCATGATTATGAGAAAGTGAAGAGCTGAGAGTGGCTGACATTTC 849
      :::::|||||
Db      260  LeuPheThrGlu-----GlnProGluGlnGlnAlaGlnArgIleIleAlaAla 275
      :::::|||||
Qy      850  TGTGTGTAACAATGCGCTCAGACTGTGAGGCCCTGTGAGTGGCTGAGGACAAACCTCC 909
      :::::|||||
Db      276  AlaGlyCysGluLysSerSerSerAlaIleValGluCysLeuArgGluLysThrGlu 295
      :::::|||||
Qy      910  AAGAGAGCTG-----CTGACCTCAGCCAGAAACAAAGCTTTTACCTCGAGTGGTAT 963
      :::::|||||
Db      296  AlaGluMetGluGlnIleThrLeuLysMetProProMetPheIleSerAlaSerLeuAsp 315
      :::::|||||
Qy      964  GGTGCTTCTTCTTCTTAATGAGCTCTAGATCTAATTTGTCCAGAAAGCTTTAAACAAT 1023
      :::::|||||
Db      316  GlyValPhePheProLysSerProArgGlnLeuSerGluLysValIleAsnAlaVal 335
      :::::|||||
Qy      1024  CTTTCATCATCGAGATCAATACCAAGAGTGGCTTCTGCTGCTCCT--ATGAAGAG 1080
      |||||:::|||||
Db      336  ProTyrIleIleGlyValAsnAsnCysGluIlePheGlyTyrIleLeuProArgMetCys 355
      :::::|||||
Qy      1081  GCTCTGAGATCCCTCAGTGGCTCCAAAGATCCCTTGGCTCCATCTGATCAAAACATC 1140
      :::::|||||
Db      356  PheProGluPheThrGluGlyLeuGluLysAspAlaIleArgGlnValIleGlnSerThr 375
      :::::|||||
Qy      1141  CTGACATC-----CCGCTCAGTATTGTCACCTTGGTGGCTTAATGATATAC 1185
      |||||:::|||||
Db      376  LeuAlaLeuSerPheLysGlyAlaPheSerAspIleValAspLeuValTyrAsnGluTyr 395
      :::::|||||
Qy      1186  TTCATATACAAAGCACTCCCTGACTGAAATCCGAGACAGCTTCTTGACTTCTTGAGAT 1245
      :::::|||||
Db      396  IleGlyValAlaGluAsnArgAlaGlnValArgAspGlyLeuLeuAspSerIleAlaAsp 415
      :::::|||||
Qy      1246  GTGTTCTTGTGGCTCCTGACTGATCAAGCTGATATACAGAGTGGCTGGTGGACCT 1305
      |||||:::|||||

```

```

Db      416  ProLeuPheValPheSerAlaValGluValAlaArgHisIleArgAspAlaGlyAsnPro 435
      |||||:::|||||
Qy      1306  GTTACTTCTATGAGTTTCCGACCCGCTCAGTGTCTTGAAGACACGAGCCGGCTTTT 1365
      :::::|||||
Db      436  ValTyrPheTyrGluPheGlnHisIleArgProSerSerAlaIleGlyValValProGluPhe 455
      :::::|||||
Qy      1366  GTCAAGCCGACACCGCTGATGAAGTCCGCTTGGTTCGCTGGTGGCTTCTTGAAGGG 1425
      |||||:::|||||
Db      456  ValLysAlaAspHisAlaAspGluIleAlaPheValPheGlyLysProPheLeuAlaGly 475
      :::::|||||
Qy      1426  GACATTTGATATGTTCAAGAGCCACGAGAGAGAGAAATTATCTGAGCCGGAATGATG 1485
      :::::|||||
Db      476  Asn-----AlaThrGlnGlnGlnAlaLysLeuSerArgThrValMet 489
      :::::|||||
Qy      1486  AAATATCTGGCTACCTTGTCTCGAACCAGGAATCTTAATGGGAAACGACTGTCTGTGG 1545
      |||||:::|||||
Db      490  LysTyrTyrPheThrAsnPheAlaArgAsnGlyAsnProAsnGlyGluGlyLeuValHis 509
      :::::|||||
Qy      1546  CGAGCTTATATCTGACTGAGACGACTCCAGCTGAGCTTGAACATGAGCTCGACAG 1605
      |||||:::|||||
Db      510  ProGlnTyrAspMetAspGluArgTyrLeuGlnIleAspLeuThrGlnLysAlaAlaLys 529
      :::::|||||
Qy      1606  AGACTCAAAAGACCGCGGTGGATTTTGTGACACGACCAATCCCCCTGATCTGTGCC 1665
      :::::|||||
Db      530  LysLeuLysGlyLysArgLysMetGluPheThrMetGlnLeuThrGlnIleMetSerAsp 549
      :::::|||||
Qy      1666  TTCGACATGCTCCACATGCTCTT 1689
      |||||:::|||||
Db      550  ArgArgArgLysLysHisThrAspLeu 557
      :::::|||||

RESULT 2
A34329
60K esterase (BC 3.1.1.-) isoform 2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 22-Jun-1990 #sequence_rev1sun 22-Jun-1990 #text_change 09-Jul-2004
C:Accession: A34329
R:Ozols, J.
J. Biol. Chem. 264, 12533-12545, 1989
A:Title: Isolation, properties, and the complete amino acid sequence of a second form of
A:Reference number: A34329; MUID:89308686; PMID:2745458
A:Accession: A34329
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-532 <O20>
A:Cross-references: UNIPROT:P14943
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
F:32-517/Domain: cholinesterase homology <CHE>
F:201,430/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 3 086-87 Length: 532
Score: 1231.50 Matches: 253
Percent Similarity: 62.19% Conservative: 76
Best Local Similarity: 47.83% Mismatches: 167
Query Match: 38.85% Indels: 33
DB: 2 Gaps: 7

US-10-023-515-3 (1-1746) x A34329 (1-532)
Qy      109  CCAACAGAGAACCAAGCTGGAGTGAATTCAGGGAGCAAGCACTGCTGCTGGAAGC 168
      |||||:::|||||
Db      6  ProIleArgAsnThrHisThrGlyGlnValArgGlySerLeuValHisValGluGluTyr 25
      :::::|||||
Qy      169  CCTGTGCTGTGAACGTGTTCCTCGAGAGTCCCTTTGGTGGTCTCCCGCTGGAGATCC 228
      |||||:::|||||
Db      26  AspAlaGlyValHisThrPheLeuGlyIleProPheAlaLysProProLeuGlyProLeu 45
      :::::|||||
Qy      229  CGATTATACGAACCGGAGCTGATCGCCCTGGAGTAACCTGTGGAGAGCAAGCACTCTAC 288
      |||||:::|||||
Db      46  ArgPheAlaProProGluProAlaGluAlaTyrPheSerGlyValArgAspGlyThrSerLeu 65
      :::::|||||
Qy      289  CTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTATGATCAACATGCTCAAGGTG 348
      |||||:::|||||

```



```

Db      ProAlaMetCysLeuGlnAsn-----LeuAlaIleMetAspGlnAspValLeuLeuLeu 83
Qy      CATTAAC-----CCGAATTCGGAGTCACAAAGATCCCTTACTTGAACATCTATGG 402
Db      HisPheThrProPheSerIlePheMetSerGlnAspCysLeuTyIleuAsnIleTyIleuSer 103
Qy      CCTGGCCCAAGCGGATACAGAGCTCCCAAGCTCCCGCTGGTGGTGGTCCCAAGAGTGGCC 462
Db      ProAlaHisAlaIleValGlnGlySerAspLeuProAlaMetValTyrIleHisGlyGly 123
Qy      TTCAAGACTGGCTCAGCTCCATCTTGAATGGCTCCGCTGGCTGGCTATGAGAGCTG 522
Db      LeuThrMetGlyMetAlaSerMetTyrAspGlySerAlaLeuAlaIlePheGlyAspVal 143
Qy      CTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 582
Db      ValValValIleThrIleGlnIleTyrArgLeuGlyValLeuGlyPhePheSerThrIleAspGln 163
Qy      CATGCTCCGGGGAATCGGCGCTTCAAGGACGAGTGGCTGTGCTGGTGGTGGTGGTGGTGG 642
Db      HisAlaThrGlyAsnHisGlyTyrIleuAspGlnValAlaAlaLeuArgTyrValGlnIle 183
Qy      AACATCGAGTCTTCTGGTGGGAGCCCAAGCTCTGTGACCATCTTGGCGAGTCCGCGGA 702
Db      AsnIleAlaHisPheGlyGlyAsnProGlyArgValThrIlePheGlyGlySerHisGly 203
Qy      GCGATAGTGTCTTGAATCTTACTGTCTGTCTCCATGGCCAAAGCTTATTCACAAAGCC 762
Db      GlyThrSerValSerSerHisValLeuSerPrometSerGlnGlyLeuPheHisGlyAla 223
Qy      ATCATGAGAGTGGGAGGCGCATCCCTTACCTGGAGGCGCCATGATTAAGAAAGAT 822
Db      IleMetGlySerLeuValAlaLeuLeuProGlyLeuIleThrSerSerSerGlyValVal 243
Qy      GAGGACCTGACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 882
Db      Ser-----ThrValValAlaAsnLeuSerArgCysGlyGlnValAlaAspSerGlnThrLeu 261
Qy      CTGAGGCTGCTGAGGACAAACCCCTCCAAAGAGCTGTCGACCTTCAAGCCAAAGAA 942
Db      ValArgCysLeuArgAlaIleSerGlnIleGlnIleMetLeuAlaIleThrGlnValPheMet 281
Qy      TCTTTCACTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1002
Db      LeuIleProGlyValValAlaAspGlyValPheLeuProArgHisPheGlnGlyLeuLeuAla 301
Qy      CAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACCAAGAGTGGCTTC 1062
Db      LeuAlaAspPheGlnProValProSerIleIleGlyIleAsnAsnAspGlnIleTyrTrp 321
Qy      CTGCTGCTGCT-----ATGAGAGAGCT 1083
Db      IleIleProIleValLeuLeuAlaIleAspProGlnGlnGlyArgAspArgGlnAlaMet 341
Qy      CCTGAGATCTTCAGAGGCTCCCAAGATCCCTTGGCTCCATCTGATACAAACATCTCG 1143
Db      ArgGlnIleMetHisGlnAlaThrIleValLeuMetLeu----- 354
Qy      CACATCCGCGCTCAGATATTGTCACCTTGGCTGAATGAATCTTCATGACAAAGATCC 1203
Db      ProProAlaLeuGlnGlyAspLeuMetAspGlnIleMetGlySerAsnGlyAsp 372
Qy      CTGACTGAATTCGGAAGAGCTTCTGCACTTGGAGATGTGTTCTTGTGGCTCT 1263
Db      ProIleValIleMetAlaGlnPheGlnGlnIleMetMetAlaAspAlaMetPheValMetPro 392
Qy      GCACTGATCACAAGCTCGATATACAGAGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1323
Db      AlaIleuArgValAlaHisIleGlnIleArg--SerHisAlaProThrTyrIlePheIle 411
Qy      CGGACCGGCGCTCAGTGGCTTGAAGACGAGCCGCGCTTGTGCAAGCCGACGACGCT 1383

```

```

Db      GlnHisArgProSerPheThrIleValAspLeuArgProPheHisValAlaAspHisGly 431
Qy      GATGAAGTCCGCTTGTGTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1443
Db      AspGlnValValPheValPheArgSerHisValLeuPheGlySerIleValProLeu----- 449
Qy      GAGGCCAGGAGAGAGAGAAATTAATGAGCCGGAAGATGATGAATTAATGAGCTT 1503
Db      ThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 467
Qy      GCTGACACCGGAGATCTTAATGGGAACGACCTGTCTGTGGCCGACCTTAATTAATGACT 1563
Db      AlaArgAsnArgAsnProAsnGlyGlyGlyLeuAlaHisIleThrProLeuPheAspLeuSer 487
Qy      GAGCAGTACCTTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1623
Db      GlnArgTyrLeuGlnLeuAsnMetGlnProAlaValGlyGlnAlaLeuValAlaArgArg 507
Qy      GTGATTTTGGACCAAGACCATCCCTCC 1650
Db      LeuGlnPheThrThrHisThrLeuPro 516

RESULT 3
JC5408
carboxylesterase (EC 3.1.1.1) - human
C.Species: Homo sapiens (man)
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C.Accession: JC5408
R.Schwer, H.; Langmann, T.; Daif, R.; Becker, A.; Aelaniadis, C.; Schmitz, G.
Biochem. Biophys. Res. Commun. 233, 117-120, 1997
A>Title: Molecular cloning and characterization of a novel putative carboxylesterase, pre
A.Reference number: JC5408, MUID:97289502, PMID:9144407
A.Accession: JC5408
A.Molecule type: mRNA
A.Residues: 1-559 <SCH>
A.Cross-references: UNIPROT:000748; GB:Y09616; MID:92058317; PIDN:CAA70831.1; PID:9205833
A.Experimental source: intestine
C.Comment: This enzyme hydrolyzes many xenobiotics, such as carboxyl esters, thioesters &
C.Genetics:
A.Gene: GDB:CEB2; ICE; CE2
A.Cross-references: GDB:9959011
C.Superfamily: Cholinesterase; cholinesterase homology
C.Keywords: carboxylic ester hydrolase; glycoprotein
F:58-544/Domain: cholinesterase homology <CHS>
F:15-95,123-280,291-428/Disulfide bonds: #status predicted
F:111,276/Binding site: carbonylate (Asn) (covalent) #status predicted
F:128,457/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 4,076-87 Length: 559
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 38.80% Indels: 24
DB: 1 Gaps: 8

US-10-023-515-3 (1-1746) x JC5408 (1-559)
Qy      GGGCGTCTGCTGAAGGCGCACAGAGAACGAGCTGGAGTGAATTCAGGGCAAGCA 150
Db      GlnGlnAspSerAlaSerProIleArgThrThrHisIleGlnValLeuGlySerLeu 45
Qy      GTCACTGTGCTGGGAAGCCCTGTGCTGGAACGTCTTCTGGAGTCCCTTGTGCTGT 210
Db      ValHisValIleValAsnAlaGlyValGlnThrPheLeuGlyIleProPheAlaVal 65
Qy      CCGCGCTGGAGTCCCTGAGATTTTACGAACCGGAGCTGATCCGCTCGGATTAATCTTG 270
Db      ProProLeuGlyProLeuArgPheAlaProGlnIleProGlnIleSerIleVal 85
Qy      CGAGAAGCCACTCTTACCTTAATTTGTGCTCCAGAAC-----TCAGAG 315
Db      ArgAspGlyThrThrHisPheProAlaMetCysLeuGlnAspLeuThrAlaValGlnSerGln 105

```

```

OY 316 TGGCTGCTTAGATCAACATGCTCAAGTGCATTAACCGAAATTGCGAGTGCAGA 375
DB 106 PheLeu-----SerGlnPheAseMetThrPheProSerXerPseMetSerGln 121
OY 376 GACTGCTCTACTGAAACATCTTATGGCGCTGCCCAAGCCGATACAGAGCTTCAAGCTCCC 435
DB 122 AspCyALeuIlyrLeuSerXerIlyrThrProAlaHiserXerIleGluIlySerAenLeuPro 141
OY 436 GTCTTGATGTGATGTTCCAGAGAGTGCCTTCAAGTGCCTGAGCTCCATCTTTATGAGG 495
DB 142 ValMetValTrrIleHisGlyAlaLeuValPheGlyMetAlaSerLeuIlyrAspGly 161
OY 496 TCCGCGCTGGCTGCTATGAGAGAGTGTGTGTGTGCTGCTCCAGTACCGGCTAGAGATA 555
DB 162 SerMetLeuAlaAlaLeuGluAenValValValIleIleGlnIlyrIleGluGlyVal 181
OY 556 TTTTGTTTCTTACCAACATGGAGTACGATGCTCCGGGAACTGGGCTTCAAGACAG 615
DB 182 LeuGlyPhePheSerThrIlyrAspIlyrValAlaThrGlyAenIlyrGlyrLeuAenGln 201
OY 616 GTGCTGCTGTGCTGCTGCTCCAGAAACATCGAATCTTGCTGGTGGGACCCGACCTT 675
DB 202 ValAlaAlaLeuAlyrTrrValIleGlnAenIleAlaHisPheGlyGlyAenProAspArg 221
OY 676 GTGACCATCTTTGGCGAGTCCGCGGAGGACATAGTGTCTTATGCTTACTGCTGCC 735
DB 222 ValThrIlePheGlyGlySerAlaGlyIlyrThrSerValSerIleuValValSerPro 241
OY 736 ATGGCCAAAGGCTTTATCCCAAAACCATGAGAGAGTGGGCTGCCATCATCCCTTAC 795
DB 242 IleserGlnGlyLeuPheHisGlyAlaIleMetGlnSerGlyValAlaLeuLeuProGly 261
OY 796 CTGAGAGCCCATGATTATGAGAAGAGTGAAGTGCAGTCTG-----CAGATGGTGCATTTTC 849
DB 262 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaAseMetLeu 277
OY 850 TGTGTAACAATGCGCTCAAGACTCTGAGCGCTGCTGAGGTGCTGAGACAAACCTTCC 909
DB 278 SerAlaCyAAspGlnValAlaAspSerGlnAlaLeuValGlyCyALeuAlyrGlySerIlyr 297
OY 910 AAGAGCTGCTGAGCTCAAGCTCAGCAAGAAACAAAGTCTTACCTGAGTGTGATGTGCT 969
DB 298 GlnGlnIleLeuAlaIleAseMetLeuPheIlyrMetIleProGlyValValAlaAspIlyrVal 317
OY 970 TTTCTTCTTAAGACCTCTAGATCTATGCTCTGAGAAAGATTTAAAGCAATCTTCC 1029
DB 318 PheLeuProArgHisPheProGlnGluLeuLeuAlaSerAlaAspPheGlnProValProSer 337
OY 1030 ATCATCGAGTCAATACCAAGAGTGTGGCTTCTGCTGCT-----ATGAAG 1077
DB 338 IleValGlyValAlaAseMetAenGlnPheGlyIlyrLeuIleProIlyrValMetArgIleIlyr 357
OY 1078 GAGGCTCTGAGATCTCAAGTGCCTCAAGACCTGCTGCTGCTGCT-----ATGAAG 1137
DB 358 AspThrGlnIlyrGlyMetAspArgGlyAlaSerGlnAlaAlaLeuGlnIlyrMetLeuThr 377
OY 1138 ATCTGCAACATCCGCGCTCAGATTTGACCTGTGAGCTTATGATTAATCTTCAACAAG 1197
DB 378 LeuLeuMetLeuProIlyrThrPheGlyAspLeuLeuArgGlnIlyrIleGlyAspAse 397
OY 1198 CAGTCCCTGAGTAAATCCGAGACAGTCTTGTGAGCTTGTGAGATGTGTTTGTG 1257
DB 398 GlyAspProGlnIlyrLeuGlnAlaGlnPheGlnIlyrMetMetAlaAspSerMetPheVal 417
OY 1258 GTCCCTGACTGATCAGAGCTCATATACAGATGCTGCTGCTGCTGCTGCTTAT 1317
DB 418 IleProAlaLeuGlnIlyrAlaHisPhe---GlnCySerArgAlaProValIlyrPheIlyr 436
OY 1318 GAGTTTGGGACCGCGCTCAGTGTGTAAGACAGACAGAGCGCTTTGTCAAGCGCAC 1377
DB 437 GlnPheGlnIlyrGlnProSerTrrPheIlyrAseMetIleArgProIlyrHisMetIlyrAlaAsp 456

```

```

OY 1378 CAGCGTGAAGATGCGGCTTGTGTGCTGCTGCTTCTGAAAGGAGCAATTGTATG 1437
DB 457 HisGlyAspGlyLeuProPheValPhe---ArgSerPhePheGlyGlyAenIlyrIleIlyr 475
OY 1438 TTGAAAGAGCCAGGAGAGGAGGATTAAGTACGCGGAAAGATATGATACTAGGCT 1497
DB 476 Phe-----ThrGlnGlnIlyrGlnIleuSerArgIlyrMetMetIlyrIlyrAla 492
OY 1498 ACCTTGTCTGAAACCGGGAATCTTAATGGAAACGACTGTCTGTGGCCAGCTTATAT 1557
DB 493 AsnPheAlaArgAenGlyAenProAenGlyGlyLeuProAlaIlyrProLeuPheAsp 512
OY 1558 CTGACTGAGCAGTACCTTCAAGTGCCTGAAACATGAGCTCCGAGACAGACTGAADA 1617
DB 513 GlnGlnGlnIlyrIlyrGlnIleuAenMetLeuGlnProAlaValGlyArgAlaLeuIlyrAla 532
OY 1618 CCGCGGTGATTTTGGACGACGACCATCC 1650
DB 533 HisArgLeuGlnPheTrrPheIlyrAlaLeuPro 543

RESULT 4
847655
carboxylesterase (EC 3.1.1.1) precursor - golden hamster
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 26-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S47655
R:Some, T.: Isobe, M.; Takahatake, E.; Wang, C.Y.
Biochim. Biophys. Acta 1207, 138-142, 1994
A>Title: Cloning and sequence analysis of a hamster liver cDNA encoding a novel putative
A:Reference number: S47655; MUID:94318655; PMID:8043605
A:Accession: S47655
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <SON>
A:Cross-references: UNIPROT:Q64419; EMBL:D28566; NID:G531238; PIDD:BA05913.1; PIR:G5312;
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
P:58-546/Domain: cholinesterase homology <CHS>
P:1227,459/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 1,03e-84 Length: 561
Score: 1199.00 Matches: 252
Percent Similarity: 61.16% Conservative: 74
Best Local Similarity: 47.28% Mismatches: 181
Query Match: 37.82% Indels: 26
DB: 2 Gaps: 7

US-10-023-515-3 (1-1746) x S47655 (1-561)
OY 91 GGGGCTTGTGCTGAAGGCGCAGAGAGACAGAGCTGGATGATTCAGGGCAAGCA 150
DB 26 GlyGlnAspSerValSerProIleArgAenThrHisThrIlyrGlnValAlaGlyIlyrLeu 45
OY 151 GTCACTGTGCTGGAAGCCCTGTGCTGAGAGTGTCTTCTGAGATCCCTTGTGCTGT 210
DB 46 ValIlyrValIlyrGlnIlyrAlaThrGlyValIlyrAlaPheLeuGlyIleProPheAlaIlyr 65
OY 211 CCCCCGTGGAGATCCCTGCGATTTACGAACCGGACGCTGCAATGCGCTGGAGTAACTTG 270
DB 66 ProProValIlyrProLeuArgPheAlaProGlnProGlnProIlyrProIlyrSerGlyAla 85
OY 271 CGAAGACGACCTCCATACCTTAATTTGTGCTCCAGACTCAGAGTGGCTGCTTAAAT 330
DB 86 ArgAspGlyThrSerGlnProAlaMetCyALeuGlnIlyrAspPheMetArgProGlnIlyr 105
OY 331 CAACACATGCTCAAGGCTAATCCGAAATTCGAGTGCAGAGTGCAGAGTGCCTTACTG 390
DB 106 SerIlyrGlnArgIlyrIleIleuProThrIleSerMetSerGlnAspCyALeuIlyrLeu 125
OY 391 AACATCTAAGCGCTGCCACGCGGATACAGGCTCCAAAGCTCCCGCTTGTGTGCTTC 450
DB 126 AsnIlyrThrProAlaHisValIlyrGlnIlyrSerAenLeuProValMetValTrrIle 145

```

```

QY 451 CCAGAGGTCCTTCAGAGTCCTCAGCTTCATCTTGATGGGCGCCCTGGCTGC 510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 146 HtGtGtGtAlaLeuValMeGtGtMetAlaSerMetAlaSerGtGtLeuValAla 165
QY 511 TATGAGAGTCGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 ThrGtGtAlaLeuValLeuValSerLeuValGtGtGtGtGtGtGtGtGtGt 185
QY 571 ACATGGGATCAGCATGCTCCGGGAACTGGGCTTCAGAGACAGTGGCTCTGTCC 630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 ThrGtGtAlaLeuValAlaGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 205
QY 631 TGGGTCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 690
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 206 TrpAlaGtGtAlaLeuValAlaSerPheGtGtGtGtGtGtGtGtGtGtGtGt 225
QY 691 GAGTCCGCGGAGCCATGATGATGATGATGATGATGATGATGATGATGATGAT 750
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 ValSerAlaGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 245
QY 751 TTCCCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 246 PheHtGtGtAlaLeuValMetGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 264
QY 811 TATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 870
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 265 ThrProGtAlaValAlaGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 284
QY 871 TCTGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 285 SerGtAlaLeuValHtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 304
QY 931 CAGAAACAAAGTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 990
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 305 GlnValPheLeuMetThrProGtAlaValAlaSerGtGtGtGtGtGtGtGt 324
QY 991 GATCATGTCCTCAGAGACATTTAAAGCAATCTTCATGATGATGATGATGAT 1050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 325 GlnLeuValAlaSerValAlaSerPheHtProValProSerLeuGtGtGtGt 344
QY 1051 GAGTGTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 345 GlnGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 364
QY 1084 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 365 ArgGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 380
QY 1144 CACATCCCGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 381 ---LeuProProGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 399
QY 1204 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 400 ProGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 419
QY 1264 GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 AlaLeuValAlaLeuValPheGtGtGtGtGtGtGtGtGtGtGtGtGtGt 438
QY 1324 CGGACCGGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 GlnHtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 458
QY 1378 CAGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 HtGtGtAlaPheValAlaPheValPheGtGtGtGtGtGtGtGtGtGtGtGt 478
QY 1432 GTTATGTCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 479 -----ThrGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 492

```

```

QY 1492 TGGGTCATCTTGTGTCGAGACCGGAGATCTTAATGGAAGACATGTCCTGTGGCCAGCT 1551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 493 TrpAlaLeuPheAlaArgHtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 512
QY 1552 TATAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 513 LeuValHtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 532
QY 1612 AAAGAACCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 533 LysSerArgGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 545

RESULT 5
S34607
Carboxylesterase (EC 3.1.1.1) - mouse
C.Species: Mus musculus (house mouse)
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C.Accession: S34607
R.Aids: K.; Moore, R.; Negishi, M.
Biochim. Biophys. Acta 1174, 72-74, 1993
A.Title: Cloning and nucleotide sequence of a novel, male-predominant carboxylesterase in
A.Reference number: S34607; MUID:93326638; PMID:7916639
A.Accession: S34607
A.Status: preliminary
A.Molecule type: mRNA
A.Residues: 1-554 <AID>
A.Cross-references: UNIPROT:O63880; GB:S64130; NID:g404388; PIDN:AAE27606.1; PID:g404389
C.Superfamily: Cholinesterase; cholinesterase homology
C.Keywords: carboxylic ester hydrolase
F.46-536/Domain: cholinesterase homology <CHS>
F.215.443/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 9.57e-84 Length: 554
Score: 1186.50 Matches: 256
Percent Similarity: 61.16% Conservative: 81
Best Local Similarity: 46.46% Mismatches: 183
Query Match: 37.43% Indels: 31
Gaps: 10

US-10-023-515-3 (1-1746) x S34607 (1-554)
QY 37 TGCCTTTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 CysLeuLeuLeuLeuPheProThrThrVal-----1LeuGlyPro 15
QY 97 TCTGTGAGGCGCACAGAGAACACAGGCTGGATGATGATGATGATGATGATGAT 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 16 LysValThrGlnProGtAlaValAlaSerThrProLeuGtGtGtGtGtGtGtGt 35
QY 157 GTGCTGGAGACCTGTGCTGTGAACTGTTCTGGAAGTCCCTTGTGCTGCTCCCG 216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36 ValLysPheThrAspArgMetValAlaPheLeuGtGtGtGtGtGtGtGtGtGtGt 55
QY 217 CTGGATCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56 LeuGlyProLeuArgPheSerAlaProLeuProGlnProGtGtGtGtGtGtGtGt 75
QY 277 GCCACTCTCACTTAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 336
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 76 AlaSerLeuLeuProPheMetCysLeuGlnAlaValGtGtGtGtGtGtGtGtGtGt 95
QY 337 ATGCTGAAGTGCATTAACCCGAATTTGGAAGTGCAGAGATGCTCTTACTGAACATC 396
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 ThrLeuAlaGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 115
QY 397 TATGCGCTGCGCCAGCGGATACAGGCTCCCAAGCTCCCGTGTGATGATGATGAT 456
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 116 TyrSerProThrGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 135
QY 457 GGTGCTTCAAGACTGCTGATGCTCTCATCTTGAATGAGTCCGCTGCTGCTGATGAG 516
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 136 GlySerLeuArgValGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGtGt 155

```

OY		517	GACGTCGTGGTGGTGCGCCAGATACCGGCTGAGAAATTGGTTCTTACCACCATGG	576
Db		156	AAPValValValThrValGlnTyPAlaGlnGlyIlePheDylPheuleSerThrngLy	175
OY		577	GATCAGCATGCTCCGGGGAATTGGGCTTCAMGACAAGGTGGCTGCTGTCCTGGGTC	636
Db		176	AAPLYNHlmeCProlgYAmdArgglyPheleudApValValAlalAleuHArgTrpVal	195
OY		637	CAGAAGAATCATGAGTCTTCCGTGGGGNACCCCAAGCTCTGTACCAATCTTTGGCSAATGC	696
Db		196	GInglYAmdlLeAlarProheglgLYaPrOAsnCyeValThrllePheglYAmdSer	215
OY		697	GGGGAGCCATAAGTGTTCAGTCTTAAGCTGTCGCCATGGCCCAAAGACTTAATCCAC	756
Db		216	AlAglyGLyIllellevalserleuLeuleuSerPromelSerAladyleuPheHls	235
OY		757	AAAGCCATCATGGAGAGTGGGGTGCCATCATCTCTTACCTGGAGGCCATGATTATGAG	816
Db		236	ArgAlalleSerGlmseryAlValVallleSertLyslleleGlu-----Aspleuan	253
OY		817	AAGAGTAGGACCTGAGGTGGTGGACAT-----TTCGTGGTAAACAACGTCAGAC	870
Db		254	AlATrpSerGlnAlaGlnmepHealaaSerValAlaCyedly---SerAlserPro	272
OY		871	TCTGAGCCCTGCTGAGGTGCTGAGGACAAAACCTCCAAAGAGCTGTGACCCCTCACG	930
Db		273	Alaglu-----LeuValGlnCysleuleuGlnTybgluglyLYaAspneuHethryLSlys	291
OY		931	CAGAAAACAAATCTTTTCACTCGAGGTGTGATGAGTGCTTTCTTCTTAATAGACTCTTA	990
Db		292	AsnValAlamilleSetryThr-----ValAAspSerPheProGlnnaGrpProGln	309
OY		991	GATCTATTGTCTCAGAAAAGCATTTAAAGCAATCTCTTCATCATCGAGTCAATTAACAC	1055
Db		310	LysleuleuAlaalasnlyseglInpheProthrValProtyrleuendclYalIthrsmHls	329
OY		1051	GAGTGGCTTCTCTCTG-----CCTATGAAAGAG--GCT	1083
Db		330	GluPheglTYrPlleuleuLeuLysbhetrPaenilleeAspLyAmecGlunHlsreusEr	349
OY		1084	CCTGAGATCTTCAGTGGCTCCAAACAAGTCCCTGGCCCTCATCTGATTCAAAACATCCG	1143
Db		350	GlnGluAspLeuLeuGlnUasnsErArgrProleuEulaHls-----Mec	364
OY		1144	CACATCCCGCTCAGATTATGACCTTGCGTGAATGAATGACTTCATGACAAAGCATCC	1203
Db		365	GlnleuProFroglnullemerCProthrVallleAspGlnuryrLeuAspandlyserAsp	384
OY		1204	CTGACTGAATCCCGACAGTCTTCTGCACTTCTTGAGAGNGTCTTCTTGGTCCCT	1263
Db		385	GlUsErAlatmrArqTyralaleuendleuLeuGlyAspIlethrleuValllePro	404
OY		1264	GCACGTGATCAGAGCTGATATCAGACAGATGCTGTGACCGCTGACTTCTATGAGTTT	1322
Db		405	ThrlenulePheSerlyetyrleuGlnmephAlaglyCySProvalPheleutyrdluhe	424
OY		1324	CGGACCGGCTCAGTCTTTGAGACACGAAGCCGCTTTGTCAAGCCGACCAACCT	1383
Db		425	GlnHlstrhrProSerSerPheAlalybPheLysProAlatrVallyAlaAsphlsSer	444
OY		1384	GATGAAGTCCGCTTGTGTGGTGGTGGTCCGCTTCG-----AAGGGGACATTTGTTAG	1437
Db		445	SerGlnuAmAlapheValPheGlydlyrPheleumthrAspGlnSerleuEuleuAla	464
OY		1438	TTGCAAGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATGAATATCTGGCT	1497
Db		465	PheProGlnAlatrnglUGlnGluTybglInleuSerleuThrMeMetalaglntprSer	484
OY		1498	ACCTTGTGTGACCCGGGAATCCCTAATGGGAACAACACGTCCTCTGTGGCCGATTATAT	1557
Db		485	GlnPheAlaahglnrglyAsnpProasmndilyybelyeuProtyrProtyrProGlnleuAsn	504

```
Oy 1558 CTGACTGAGCAGTACCTCCACAGCTGACCTTGAAACCTGAGCCCTGGACAGAGACTCAAAABA 1611
      |||:::|||||:::|||||:::|||||:::
Db 505 GlnleuglunlntyrleuglnllecglyleuclurproargthrglyValylsleuylslys 524
      |||:::|||||:::|||||:::|||||:::
Oy 1618 CCGGGGGTGAATTTTGGACCAAGACCAATCCCC 1650
      |||:::|||||:::|||||:::|||||:::
Db 525 GlyArgleuGlnpnetrrPthnGlnurthneupro 535
      |||:::|||||:::|||||:::|||||:::
```

RESULT 6

carboxylesterase (EC 3.1.1.1) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 09-Jul-2004
C:Accession: A39060
R:Ovitic, M.; Tepperman, K.; Medda, S.; Elliott, R.W.; Stephenson, D.A.; Grant, S.G.; Ganer
Genomics 9, 344-354, 1991
A:Title: Characterization of a murine cDNA encoding a member of the carboxylesterase multigene
A:Reference number: A39060; MUID:91169540; PMID:1840565
A:Accession: A39060
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <OVA>
A:Cross-references: UNIPROT:P23953; GB:M57960; NID:9192853; PIDN:AAA63297.1; PID:9192854
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase
E:50-540:Domain: cholinesterase homology <CH>
E:221,455:Active site: Ser, His #status predicted

46 CTGATTCTCCAGCCCTCTGTGGGACACAGACAGTGGGGAAAAAACTGGGCTTCTGTGGAA 105
 11 LeuAlaValCysProIleuEnglyHis-----SerLeu 22
 106 GGGCCACAGAGGACCAACGAGCTGGATGATTAAGGCAAGCAAGTCACTGTCTGGGA 165
 23 ProProValValAspThrThrgInglyValValLeuEnglySerYrIleSerLeuEngly 42
 166 AGCCCTGTGCTCTGTAAAGTGTCTCTCGGAATCCCCCTTGTGCTCCCCGCTGGGATCC 225
 43 PheGInGInProValAlaValAlaPheLeuEnglyAlaProPheAlaYsProProLeuEnglySer 62
 226 CTGCGATTATTACGAAACCCGACGCTGTCAATCGCCCTGGGATTAACCTTGACAGAAAGCCACTCC 285
 63 LeuArgPheAlaProProGInProAlaGInProIlePheSerPheVallyValAlaIleThrSer 82
 286 TACCCCTAATTGTGTGCTCCAGAACTCAGAGTGG-----CTGCTCTTAAGTCAACACATG 339
 83 TyrProPheMetCysSerGInAspAlaGlyTyrAlaYsIleLeuSerAspMetPheSer 102
 340 CTCACAGTGCATTAACCCGAAATTCGGAAGTGCACAAAGATGCCCTTACCTGAAACATCTAT 399
 103 ThrGInlyGInIleuProLeuYsIleSerGInAspCysLeuYrIleuAlaIleYr 122
 400 GCGCTGTCCACGCCGATACAGGCTCCAGTCCCGCTTGTGTGTGTTCCAGAGAGT 459
 123 SerProAlaAspLeuThrlySerSerGInLeuProValMetValTyrIleIleIleGly 142
 460 GCGTTCAAGACTGCTGAGCTCATCTGTGATGGGCGCCCGCTGGCTGTATGAGGAC 519
 143 GlyLeuValIleGlyGlyArgSerProYrAsnGlyLeuAlaLeuSerAlaHisIleGluAsn 162
 520 GTGCTGTGTGTGTGTCGTCACGATACCGGCTAGGAATATTGGTCTTTCACCAATGAGAT 579
 163 ValValValAlaIleTrpIleGInlyArgLeuGlyYleIleTrpIleuPheSerThrGlyAsp 182

```
OY 580 CAGATGCTCCGGGAACTGGGCTTCAAGACGAGTGGCTGTCTGTCGTCGGGCGG 639
DB 183 GlnHisSerProGlyAsnThrPalaHisLeuAspGlnLeuAlaHisLeuArgTrpValGln 202
OY 640 AAGAACATGAGTTCTTCGTGGGAGACCCAGCTGTGACCATCTTGGCCGAGTCCGG 699
DB 203 AsphenilLeaAsnPhagGlyLysAsnProAspSerValThrIlePheGlyLysSer 222
OY 700 GGAGCCATTAAGTTTCTTAAGTTTACTGTCTCCCAAGGCCAAAGGCTTATTCCAA 759
DB 223 GlyGlyIleSerValSerValLeuValLeuSerProLeuGlyLysAspLeuPheHisArg 242
OY 760 GCCATCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 819
DB 243 AlaIleSerGlnSerGlyValValIleAsnThrAsnValGlyLysAsnIleGlnAla 262
OY 820 AGTGAGACCTGACAGTGGTGGACATTTCTGTGGTAACAATGCGTCAAGCTGAGGCC 879
DB 263 ValAsnGlnIleIleAlaThrLeuSerGlnCys-----AsnAspThrSerSerAla 280
OY 880 CTGCTGAGAGTGGCTGAGACAAACCTCCAAAGAGCTGACCTCAGC----- 930
DB 281 MetValGlnCysLeuArgGlnLeuArgGlnLeuSerGlnLeuGlnIleSerGlyLysLeu 300
OY 931 ---CAGAAACAAAGTCTTTCATCTGAGAGTGTGAGTGTGCTTCTTCTTAATAG 987
DB 301 ValGlnTyraAsnIleSerLeuSerThrMetIleAspGlyValValLeuProLysAlaPro 320
OY 988 CTAGATCATTTGTCTCAGAAAGCATTTAAAGCATTTCTTCATCATGAGTCAATAC 1047
DB 321 GlnGlnIleLeuAlaGlnLysSerPheAsnThrValProTyrlleValGlyPheAsnLys 340
OY 1048 CACGAGTGTGCTTCTCTGCTGCTATG-----AAGAGGCTCTGAG--ATCTC 1095
DB 341 GlnGlnPheGlyTrpIleIleProMetCLeuGlnAsnLeuLeuProGlnLysLysMet 360
OY 1096 AGTGCTCCAAACAAGTCCCTGCTCCATCATGATCAAAACATCTGACATCCCGCT 1155
DB 361 AsnGlnGlnThrAlaSerLeuLeuLeuArgArgPheHisSerGlnLeuAsnIleSerGln 380
OY 1156 CAGTATTTGACCTTGTGGCTAATGATATCTTCATGACAGACACTCCCTGACGAAATC 1215
DB 381 SerMetIleProAlaValIleGlnGlnTyraLeuArgGlyValAspAspProAlaLysLys 400
OY 1216 CGAGACAGTCTTCTGAGCTTGTGAGATGTGTTCTTGTGTCCTGACCTGACATCACA 1275
DB 401 SerGlnLeuIleLeuAspMetPheGlyAspIlePhePheGlyIleProAlaValLeuLeu 420
OY 1276 GCTGATATCAACAGATGCTGTGACCTGTCTACTTCTTAAGATTTGGGACCGGCT 1335
DB 421 SerArgSerLeuArgAspAlaGlyValSerThrTyraMetTyraGlnPheArgTyraPro 440
OY 1336 CAGTCTTTGAAGACAGACAGCGGCTTTTGTCAAGCGGACAGCTGATGAATCGGC 1395
DB 441 SerPheValSerAspLysArgProGlnThrValGlnGlyAspHisGlyAspGlnIlePhe 460
OY 1396 TTTGTGTTGGTGGTGGCTTCTCTGAGAGGAGGACATTTGATTTGTAAGAGCCAGG 1455
DB 461 PheValPheGlyAlaProLeuLeuLys-----GlnGlyLysSerGln 474
OY 1456 GAGGAGAAAGTTACTGAGCCGGAAGATGATGAATATCTGGGCTACTTTGCTGAAACGG 1515
DB 475 GlnGlnThrAsnLeuSerLysMetValMetLysPheTrpAlaAsnPheHisArgAsnGly 494
OY 1516 AATCTTAATGGAAGACAGCTGTCTCTGAGGACAGCTTAAATCTGAGTCAAGACACTC 1575
DB 495 AsnProAsnGlyGlnGlyLeuProHisTrpProGlnTyraAspGlnGlnGlnGlyLysLeu 514
OY 1576 CAGCTGACCTTGAAACATGAGCTCGGACAGAGACTCAAGAAACCGCGGGTGAATTTTGG 1635
DB 515 GlnIleGlyAlaThrThrGlnGlnAlaGlnArgLeuLysAlaGlnGlyValAlaLysPheTrp 534

OY 1636 ACCAGCACCATC 1647
DB 535 ThrGlnLeuLeu 538

RESULT 7
S19307
N:Altermate names: (EC 3.1.1.1) precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
J:Accession: S19307; S23607
R:Matushima, M.; Inoue, H.; Ichinose, M.; Tsukada, S.; Miki, K.; Takahashi
FBS Lett. 293, 37-41, 1991
A:Title: The nucleotide and deduced amino acid sequences of porcine liver proline-beta-n
A:Reference number: S19307; MUID:92070571; PMID:1959668
A:Accession: S19307
A:Molecule type: mRNA
A:Residues: 1-566 <MAT1>
A:Cross-references: UNIPROT:Q29550; EMBL:X63323; NID:g1930; PIDN:CAA44929.1; PID:g1931
A:Accession: S23607
A:Molecule type: protein
A:Residues: 19-40 <MAT2>
A>Note: 28-Lys and 33-Leu were also found
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F:19-566/Product: carboxylesterase #status experimental <MAT>
F:51-552/Domain: cholinesterase homology <CHS>
F:180/Binding site: carbonyl hydrate (Asn) (covalent) #status predicted
F:222,467/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 6,98e-82 Length: 566
Score: 1162.50 Matches: 250
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 44.33% Mismatches: 202
Query Match: 36.67% Indels: 41
DB: 2 Gaps: 8

US-10-023-515-3 (1-1746) x S19307 (1-566)

OY 34 TGTGCTTTTCTCGATTCCTCAGCCCTGTGGGACACAGACAGTGGGAAAACTGGG 93
DB 2 TrpLeuLeuProLeuValLeuThrSerLeuAlaThrTrp-----AlaGly 19
OY 94 CTTTGTCTGAAGGCGCACAGAGAACACAGGCTGGATTCAGGCGCAAGACATC 153
DB 20 GlnProAlaSerProProValValAspThrAlaGlnGlyArgValLeuGlyLysTyra 39
OY 154 ACTGTCGTGGAGAGCCCTGTGCTGTGAACGTTCTCGAGTCCCTTGGCTGCTGCC 213
DB 40 SerLeuGlnGlyLeuAlaGlnProValAlaValPheLeuGlyValProPheAlaLysPro 59
OY 214 CCGCTGGAGTCCCTGCGATTTTACGAACCGGACGCTGATCGCCCTGGGATTAATCGCA 273
DB 60 ProLeuGlySerLeuArgPheAlaProProGlnProAlaGlnProTrpSerPheValLys 79
OY 274 GAAGCACCCTCTACCTTAATTGTGCTCCGAAC-----TCA 312
DB 80 AsnThrThrSerTyraProProMetCysCysGlnAspProValValGlnGlnMetThrSer 99
OY 313 CAGTGGCTGCTTGTAGATCAACACATGCTCAAGGTGCAATTACCGGAATTCGAGTCA 372
DB 100 AspLeuPheThrAsnGlyLysGlnArgLeuThrLeuGlnPhe-----Ser 114
OY 373 GAAGACTGCTTACTGTAACATCTATGCGCTGCGCCAGCGGATACAGGCTCAAGCTC 432
DB 115 GlnAspCybleuTyraLeuAsnIleTyraProAlaAspLeuThrLysArgGlyArgLeu 134
OY 433 CCGTCTTGTGTGTTTCCAGAGAGTGCCTTCAAGACTGCTCAGCTCCATCTTTGAT 492
DB 135 ProValMetValTrpIleHisGlyGlyLeuValLeuGlyGlyAlaProMetTyraPhe 154
```

QY 493 GGGGCGCGCGCTGGCTGGCTTGAAGACAGCTGCTGTTGGTCTCCAGTACCGGCTAGGA 552
 Db 155 G1ValVal1Leu1Ala1Asn1Ser1Val1Val1Ala1Leu1Gln1Tyr1Gly1Leu1Gly 174
 QY 553 ATATTGGTTCTTCAACAATGGATCAGACATGCTCCGGGAACTGGGCTTCAAGAC 612
 Db 175 I1etpG1y1hep1he1ser1th1gl1AerG1ln1S1er1xg11y1an1tr1pG1y1n1S1e1u1ar 194
 QY 613 CAGTGGCTCTGTCTCTGGCTCCGAGAACATCGATGTTCTTCGGTGGGACCCGAC 672
 Db 195 G1nVal1Ala1Ala1Leu1n1Str1Val1G1nG1u1Aan11Le1Aan1heG1y1A1ar1p1roG1y 214
 QY 673 TCTGTACCACTTTTGGCGAGTCCGGCGAGGACATGAATGTTTCTGTCTTAAGTCTG 732
 Db 215 SerVal1Thr1Le1pheG1y1u1Ser1AlaG1y1G1u1SerVal1SerVal1LeuVal1Leu1Ser 234
 QY 733 CCCATGGCCAAAGCTTATTTCACAAAGCCATCATGAGAGTGGGGTGCATCATCCCT 792
 Db 235 Pro1eua1Ala1y1an1Leu1n1Ser1Ala1Le1SerG1u1SerG1y1Ala1Leu1ThrVal 254
 QY 793 TACCTGAGGCGCATGATTATGAGAGAGAGAGACCTGACGCTGGTTCACATTTCTGT 852
 Db 255 A1a1e1u1Val1Arg1y1A1ar1p1e1u1y1Ala1Ala1-----A1a1y1o1G1n1Le1A1a1Leu1Ala 272
 QY 853 GGTAACAATGCGTCAGACTTGAAGCCCTGTGAGGTGCTGAGTGCCTAGACAAACCTCCAG 912
 Db 273 G1y1Cy1e1y1u1Th1r1Th1r1Se1r1Ala1Val1Phe1Valn1G1y1e1u1n1G1ln1y1S1erG1u1A1ar 292
 QY 913 GAGCTGTCACCTCCAGCCAAACAAAGTCTTTACT- 951
 Db 293 G1u1Leu1e1u1A1ar1p1e1u1Th1r1Leu1y1Met1y1S1r1he1u1Th1Leu1A1ar1Phe1n1S1e1y1A1arG1ln 312
 QY 952 -----CGAGTGGTGAATGATGCTTTCTTCTTCAATAGACCT 987
 Db 313 ArgG1u1Ser1n1A1r1p1he1u1r1p1ro1Th1r1Val1A1erG1y1Val1Leu1e1u1r1p1ro1y1u1Met1p1ro 332
 QY 988 CTAGATCTATTGCTCTAGAAAGCATTTAAAGCAATTCCTTCCATCATCGAGTCAATAAC 1047
 Db 333 G1u1G1u11Le1u1A1aG1u1y1A1ar1Phe1an1Th1r1Val1P1ro1y1r11e1Val1G1y11Le1A1n1y1S 352
 QY 1048 CACGAGTGGGCTCTCTGTCGCTTAAAGAGAGAGGCTCGTAGAATCCCTCAGTGGCTCC--- 1104
 Db 353 G1nG1u1n1p1heG1y1T1r1p1e1u1e1u1r1p1ro1Th1r1Met1e1G1y1Phe1r1o1e1u1S1erG1u1y1Leu1e1u 372
 QY 1105 AACAAAGTCCCTGGCCCTCATGTATACAAC-----ATCTGCACATCCCGCT 1155
 Db 373 A1erG1n1y1e1Th1Ala1Th1S1er1e1u1n1T1r1y1S1er1y1P1ro11e1Ala1An11e1P1roG1u 392
 QY 1156 CAGTATTGGACCTTGCTGGCTAATGAATCTTCCATGACAAACACTCCCTAGTAAATC 1215
 Db 393 G1u1Leu1Th1r1P1ro1Val1Ala1Th1A1r1A1r1y1u1y1T1e1u1G1y1G1y1Th1A1r1A1r1P1ro1Val1y1e1u1S 412
 QY 1216 CGAGCAGCTTTTGGCACTTGTGGTGAAGATGTGTTCTTTGGTCCCTGCATGATCAACA 1275
 Db 413 1y1A1ar1Leu1r1he1Leu1A1ar1p1e1u1MetG1y1A1erVal1Ala1PheG1y1Ala1P1ro1SerVal1Th1rVal 432
 QY 1276 GCTGCATATCAAGAGATCTGGTGCACACTGCTACTTCTATGAATTTCCGACACCGGCT 1335
 Db 433 A1a1A1arG1ln1n1S1A1r1A1e1y1A1a1P1ro1Th1r1y1u1Met1y1G1u1n1Phe1G1n1T1r1p1ro 452
 QY 1336 CAGTGCCTTGAAGACAGACAGCGCGCTTTTGTCAAGGCGACACAGCTGATGAATGCTCCG 1395
 Db 453 SerPhe1Ser1Ser1A1r1y1e1u1A1r1p1ro1y1n1Th1rVal11eG1y1A1ar1n1A1e1y1A1arG1u1Le1he 472
 QY 1396 TTTGTGTTCCGTGGTGGCTTCTGTAAAGGGGACATTTGATTGTTGAAAGAGCCAGGAG 1455
 Db 473 SerVal1PheG1y1Phe1r1o1Leu1e1u1y1S1e1y1A1ar-----Ala1rG1o1u 486
 QY 1456 GAGGAGAAATTAATGAGCGGAAAGATGAATAATCTGGGGCTACTTTGCTGCGAAGCGG 1515
 Db 487 G1u1G1u1Val1Ser1Leu1Ser1y1Th1r1Val1e1u1y1Phe1T1r1Ala1n1Phe1Ala1n1A1r1G1SerG1y 506
 QY 1516 AATCTTAATGGAGACACTGTCTCTGTGGCGAGCTTAATCTGACTAGACAGTACTC 1575

```

Db          507 AsnProhanglygluyluleuProHletrPrrometYrAspGlnGluGluYrleu 526
              ||||| ||| ||||| |||||
Oy          1576 CAGCGACTGGAACATGAGCGCTCGACAGAGACTCAAGAACCGCGGATGATTTGG 1635
              ||||| ||||| ||||| |||||
Db          527 GlnIleGlyValAsnThrGlnAlaAlaLysArgGluLeuylgluValAlaAphetrp 546
              ||||| ||||| ||||| |||||
Oy          1636 ACCAGCACCATC 1647
              ::
Db          547 AsnAspLeuLeu 550
              ::
RESULT 8
XK0054
carboxylesterase (EC 3.1.1.1) E1 precursor, minor form - rat
C|Species: Rattus norvegicus (Norway rat)
C|Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Aug-1998
C|Accession: XK0054
R|Takagi, Y.; Morhaasbi, K.; Kawabata, S.; Go, M.; Omura, T.
V|Biochem. 104, 801-806, 1988
A|Title: Molecular cloning and nucleotide sequence of cDNA of microsomal carboxylesterase
A|Reference number: XK0054; MUID:89174514; PMID:3235453
A|Accession: XK0054
A|Molecule type: mRNA
A|Residues: 1-549 <TAK>
A|Experimental source: liver
C|Superfamily: cholinesterase; cholinesterase homology
C|Keywords: carboxylic ester hydrolase; glycoprotein; microsome
F|1-18/Domain: signal sequence #status predicted <SIG>
F|19-549/Product: carboxylesterase E1 #status predicted <MAT>
F|50-538/Domain: cholinesterase homology <CHE>
F|729,274,302,375,476/Binding site: carboxylate (asn) (covalent) #status predicted
F|722,453/Active site: Ser, His #status predicted

```

Alignment Scores:					
Pred. No.:	1,188e-80	Length:	549		
Score:	1144.00	Matches:	238		
Percent Similarity:	60.33%	Conservative:	92		
Best Local Similarity:	43.51%	Mismatches:	183		
Query Match:	36.09%	Indels:	34		
DB:	2	Gaps:	9		

US-10-023-515-3 (1-1746) x JX0054 (1-549)	
OY	46 CTGATTCTCCAGCCCCCTGTGGACACAGATGGGGAAAAACTGGGCCCTTCTGTGA 105
Dd	11 LeuAlValcysProIleTrpGlyHis-----ProSerSerPro 23
OY	106 GGGCCCAAGAAGAACCCAGCGTGGAATTCACGGCAGACGAATCATCTGCTCTGGA 165
Dd	24 ---ProValAlaSerThrThrylselgylvalLeugllySTyValserLeuglly 42
OY	166 AGCCCTGNGCCTGTGAAGTGTTCCTCGAAGTCCCTTGCCTGCCCGCTGGATCC 225
Dd	43 PheThrGlnProValAlaValPheLeuglyalProPheAlalyBProProLeuglySer 62
OY	226 CTGCATTATACAACC CGCAGCTGCATCGCCCTGGGATAACTTGGCGAAAGCACTCC 285
Dd	63 LeuArgPheAlaProProGluProAlaGluProTrpSerPheValLyAsnThrThrThr 82
OY	286 TACCCTAATTTGTGCTCCAGAAC-----TCAGAGTGGCTCTC 324
Dd	83 TyrProPrometCysSerGlnAspelyalValglybLeuLeuAlaMetLeuSer 102
OY	325 TTATGATCAACATCCTCAAGGTGCATTACCGGAATTCGAGTGTCAAGACCTGCTC 384
Dd	103 ThrGlyLySgluSerIleProLeuGluPhe-----SerGluAspCysLeu 117
OY	385 TCCTGAACATATAGAGCGCTGCCCACGCCGATACAGGCTCCAAGCTCCCGCTTGGTG 444
Dd	118 TyzLeuAsnIleTySerProAlaAspleuthThrylasnsrtaGluLeuProValMetVal 137
OY	445 TGAGTCCAGAGGTCCTCAAGATCGCTCACGCTCCACTTGTGAATGGAGTCGGCCTG 504


```

Db      138  TrrllelsglyglyleullellellyglyalaserProtyrserglyleuallaleu 157
Oy      505  GGTGCTATATGAGAGCGTGTGGTGGTCGTCGATGCGCGGTAGAAATATGTTTC 564
Db      158  SerlAhlsGluAsnValValValThrllleGlnlyrAargleGlylTerpGlyleu 177
Oy      565  TTCAACAATGAGATCAGATGCTCCGGGAACCTGGGCTTCAGAACAGAGTGGCTCT 624
Db      178  PheSerThrGlyAspGluHleSerArglyAsnThrAlahleuAspGluLeuAlah 197
Oy      625  CTGTCTGGGTCCAGAGAATCATGAGTTCTTGGTGGGACCCCAAGCTCTGTGACATC 684
Db      198  LeuArgTrpValGlnAspAsnllleAlaAsnPheGlyGlyAsnProAspSerValThrll 217
Oy      685  TTGGGCGAGTCCGGGGAGCAATAGTGTCTTACTCTTATATGCTGTCCATGGCCAA 744
Db      218  PheGlyGluSerAlahGlyValSerValSerAlahleuValLeuSerProLeuAlah 237
Oy      745  GGTATATCCACAAAGCCATCATGAGAGTGGGTGGCCATCATCCCTTACCTGAGGCC 804
Db      238  AsnleuPheHlsArgAlahlelSerGluSerGlyValValleuThrThrAsnleuAsp 257
Oy      805  CATGATTAATGAGAGAGAGAGAGCTGAGGTGTTGCAATTCTGTGTGAACAATGCG 864
Db      258  LysAsnThrGlnAlahVal-----AlahlnMetllleAlahThrLeuSerGlyCyAsnAsn 275
Oy      865  TCAGACTGTAGGCGCTGAGTGGCTGAGGAGCAAAACCTCCAGAGAGCTGTGACC 924
Db      276  ThrSerSerAlahleuValGlnCyAsnleuArgGlnlyrThrGlnAlahleuLeuGln 295
Oy      925  CTCAGCCGAGAA-----ACAAAGCTTTCACTCGAGTGGTGTGATGCTTCTTCTT 978
Db      296  LeuThrVallylSleuAspAsnThrSerMetSerThrValllleAspGlyValleuPro 315
Oy      979  AATGAGCTCATAGATCTATGTCTTCAGAAAGCAATTAAAGCAATCTTCATCATCGCA 1038
Db      316  LysThrProGluGluGluLeuThrGlnlyrLysSerPheAsnThrValProtyrValGly 335
Oy      1039  GTCAATAACACAGATGCTGCTCTGCTGCTTATGAAGAGAGCTCCGAGATCTCACT 1098
Db      336  PheAsnlyrGlnGluPheGlylTrpIlelleProThrMetMetGlyAsnleuSerGlu 355
Oy      1099  GGC---TCACACAGTCCCTGCGCTCCATCATGATACAAACATC-----CTGCAC 1146
Db      356  GlyArgMetAsnGluLysMetAlahSerSerPheLeuLysArgPheSerProAsnleuAsn 375
Oy      1147  ATCCCGCTCAGATTTGTCACCTGTGCTATATGAATCTTCATGACAGACATCCCTG 1206
Db      376  lIleSerGluSerValllleProAlahlelleGluLyslyrLeuArgGlylThrAspAspPro 395
Oy      1207  ACTGAATCCAGACAGATCTTGTGACCTTGTGAGAGATGTTTGTGTGCTCCCTGCA 1266
Db      396  AlahlyrlyrlyrGluLeuLeuAspMetPheSerAspValPhePheGlylLeProAla 415
Oy      1267  CTGATCAGACAGCTCATCATCAGAGATGCTGGTGACCTGTCTTATCTGATGATTTGG 1326
Db      416  ValleuMetSerArgSerleuArgAspAlahGlyAlahProThrlyrMetCysGluHleGln 435
Oy      1327  CACCGGCTCAGTGTCTTGAAGACACGAGCGGCTTTTGTCAAGCCGACCACTGAT 1386
Db      436  TyrArgProSerPheValSerArgPlnArgProGlnThrValGlnGlylAspHlsGlyAsp 455
Oy      1387  GAAATCGGCTTGTGTCTGCTGGTGGTCCCTTGAAGGGGGAGCATTTATGTTGCAAGA 1446
Db      456  GluIlePheSerValPheGlylThrProPheLeuys-----GluGly 469
Oy      1447  GCCACGAGGAGGAGAGATATGATGAGGAGATGAATATCTGGGCTACTTTTGT 1506
Db      470  AlahSerGluGluGlnThrAsnleuSerlyrleuValMetlyrPheTrpAlahAsnPheAla 489
Oy      1507  CGAACCGGGAATCTTAATGGAGACACACTGTCTCTGTGGCCAGCTTAATATGACTGAG 1566
Db      490  ArgAsnGlyAsnProAsnGlyGluGlyleuPronHlsTrpProGluLysAspGlnlyrGln 509

```

```

Oy      1567  CAGTACTCCAGCTGAGTGAATGACATGACCTCGGACAGACTCAAGAACCGCGGCTG 1626
Db      510  GlyTyrLeuGlnHlleGlylAlahThrThrGlnGlnAlahGlnlyrleuLysGlyGluGluVal 529
Oy      1627  GATTTTGGACGACGACCATC 1647
Db      530  AlahPheTrpThrGluLeu 536

RESULT 9
A1010
N:Aldehyde/esterase (EC 3.1.1.1) precursor, monocyte/macrophage [validated] - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
R:Accession: A1010, JH0327, K47376, A49816, PS0280, 161085, A48809, 157004
J:Biol. Chem. 266, 18832-18838, 1991
A:Title: A serine esterase released by human alveolar macrophages is closely related to )
A:Reference number: A41010, MUID:92011649, PMID:1918003
A:Accession: A41010
A:Molecule type: mRNA
A:Residues: 1-567 <MUN>
A:Cross-references: UNIPROT:P23141, GB:W73499, NID:G179927, PIDN:AAA5649.1, PID:G179928
A:Note: parts of this sequence, including the amino end of the mature protein, were confi
R:Long, R.M.; Calabrese, M.R.; Martin, B.M.; Pohl, L.R.
Life Sci. 48, PL43-PL49, 1991
A:Title: Cloning and sequencing of a human liver carboxylesterase isoenzyme.
A:Reference number: JH0327, MUID:91148424, PMID:1997784
A:Accession: JH0327
A:Molecule type: mRNA
A:Residues: 61-567 <LON>
A:Cross-references: GB:M55509, NID:G179929, PIDN:AAA5650.1, PID:G179930
R:Shibata, F.; Takagi, Y.; Kitajima, M.; Kuroda, T.; Omura, T.
Genomics 17, 76-82, 1993
A:Title: Molecular cloning and characterization of a human carboxylesterase gene.
A:Reference number: A47376, MUID:94010913, PMID:8406473
A:Accession: A47376
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA; DNA
A:Residues: 1-3, 'PALV', 8-11, 'A', 13-567 <SHI>
A:Cross-references: GB:D21089, NID:G455476, PIDN:BA04650.1, PID:9458470
R:Note: sequence extracted from NCBI backbone (NCBI:137630) and corrected to correspond
R:Zschunke, F.; Salmassi, A.; Kneipe, H.; Buck, F.; Parwaresch, M.R.; Radzun, H.J.
Blood 78, 506-512, 1991
A:Title: cDNA cloning and characterization of human monocyte/macrophage serine esterase-1
A:Reference number: A49816, MUID:91300111, PMID:2070086
A:Accession: A49816
A:Molecule type: mRNA
A:Residues: 'G', 65-185, 'G', 187-361, 363-567 <ZSC>
A:Cross-references: GB:X52973, NID:G36421, PIDN:CAA7147.1, PID:G1335304
R:Ridides, P.W.; Richard, L.J.; Bowles, M.R.; Pond, S.M.
Gene 108, 289-292, 1991
A:Title: Cloning and analysis of a cDNA encoding a human liver carboxylesterase.
A:Reference number: PS0280, MUID:92084150, PMID:1748313
A:Accession: PS0280
A:Molecule type: mRNA
A:Residues: 114, 'H', 116-280, 'A', 282-300, 'IGNSLYLTYRTQREST', 318-336, 'R', 338-382, 'GSPT', 31
A:Cross-references: GB:M65261, NID:G187028, PIDN:AAA83932.1, PID:G187029
A:Experimental source: liver
A:Note: differences between this sequence and other reports appear to be due to frame shift
R:Kroetz, D.L.; McBride, O.W.; Gonzalez, F.J.
Biochemistry 32, 11606-11617, 1993
A:Title: Glycosylation-dependent activity of baculovirus-expressed human liver carboxylester
A:Reference number: A48809, MUID:94032283, PMID:8218228
A:Accession: 161085
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-55, 'G', 57-361, 363-535, 'G', 537-567 <KRO1>
A:Cross-references: GB:L07765, NID:G180949, PIDN:AAA35711.1, PID:G180950
A:Accession: A48809
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```

A:Residues: 1-17, 'A', 18-55, 'G', 57-535, 'G', 537-567 <R02>
A:Cross-references: GB:L07764; NID:q180947; PIDN:AA16036.1; PID:q180948
C:Genetics:
A:Gene: GDB:CE51; HMSE
A:Cross-references: GDB:128044; OMIM:114835
A:Map position: 16q13-16q22.1
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; liver
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-145/Domain: propeptide #status predicted <PRO>
F:150-553/Domain: cholinesterase homology <CHE>
F:146-567/Product: carboxylesterase #status experimental <MAT>
F:564-567/Region: endoplasmic reticulum retention signal #status atypical
F:221,468/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.:	3,88e-80	Length:	567
Score:	1140.00	Matches:	248
Percent Similarity:	58.39%	Conservative:	79
Best Local Similarity:	44.29%	Mismatches:	197
Query Match:	35.96%	Indels:	36
DB:	1	Gaps:	9

US-10-023-515-3 (1-1746) x A41010 (1-567)

```
QY 34 TGGTCTTTTTCCTGATTCTCCAGCCCTGTGGAGACACAGAGGGGAAAACTGGG 93
DB 2 TTPLEUAGLALPHALIELEUAlALHREUSERAlASERAlALALTRGlyHis----- 19
QY 94 CTTTCTGTGAAGGGCCACAGAGAACACAGCTGGATGTATTCAGGCAAGCAAGTC 153
DB 20 ProSerSerPro---ProValValAspThrValHisGlyLysValLeuGlyLysPheVal 38
QY 154 ACTGTGCTGGAGAGCCGTGCTGTGAACGTCTTCCTGGAGTCCCTTGTGCTGCTCC 213
DB 39 SerLeuGlyLysPheAlGlnProValAlAlPheLeuGlyLysPheAlAlAspPro 58
QY 214 CCGTGTGATCCCTGCGATTTACGAACCCGAGCTGCGATCGCCCTGGATTAATTCGCA 273
DB 59 ProLeuGlyProLeuArgPheThrProGlnProAlAGluProTrpSerPheValLys 78
QY 274 GAAGCCACTCTCACTCAATTTGCTGCCAAGACTGAG-----TGCTGCTCTTA 327
DB 79 AsnAlaThrSerLysProPrometCysThrGlnAspProLysAlAGlyGlnLeuSer 98
QY 328 GATCAACACATGTCACAGTGATTAACCGAAATTCCGAGTGTGAGAGATGCTCTAC 387
DB 99 GluLeuPheThrAsnArgLysGlnAsnAlLeProLeuLysLeuSerGluAspCysLeuTrp 118
QY 388 CTGAACATCTATGCGCCCTGCCACGCGATACAGGCTTCCAGCTCCCGTCTGTGTGG 447
DB 119 LeuAsnLysLysTrpProAlAspLeuThrLysLysAsnArgLeuProValMetValTrp 138
QY 448 TTCCAGAGAGTGTGCTTCAAGATGCGCTCAAGCTTCCATCTTATGGTCCGCTGCT 507
DB 139 IleHisGlyLysGlyLeuMetValGlyAlAlAsnThrLysAspGlyLeuAlAlAlAl 158
QY 508 GCGTATGAGACGTGCTGTGATGCTGATCCAGTACCGCTAGAGAAATTTGTTGTTCTC 567
DB 159 AlaHisGlnMetValValValValThrLysGlnLysArgLysGlyLysPhePhe 178
QY 568 ACCATGAGGATGATGATGCTCCGGGAACTGGGCTTCAAGACCAAGTGGCTGCTCG 627
DB 179 SerThrGlyAspGlyLysLysSerArgLysAsnTrpGlyHisLeuAspGlnValAlAlAl 198
QY 628 TCCGGGTTCAGAGACATGAGATTTCTGCTGGAGACCCCACTCTGTGACATCTTT 687
DB 199 ArgTrpAlaGlnAspAsnAlAlAlSerPheGlyLysAsnProGlySerValThrLysPhe 218
QY 688 GCGAGTCCGCGAGGACATAGTGTCTTATGCTTATGCTCTCCATGCGCAAGGC 747
DB 219 GlyLysSerLysGlyLysGlyLysSerValSerValLeuValLeuSerProLeuAlAlLysAsn 238
```

```
QY 748 TTAATTCACAAAGCATATGAGATGGTGGGCTCATCTTACTGTGAGCCCAT 807
DB 239 LeuPheHisArgAlAlLysSerLysSerGlyValAlAlLeuThrSerValLeuValLysLys 258
QY 808 GATTATGAGAAAG-----AGTGAAGACCTGCGAGCTGTGTGACATTTCTGTGTAACAT 861
DB 259 GlyAspValLysProLeuAlAGlnGlnAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 275
QY 862 GCGTCAGACTGTGAGCCCTGCTGAGTGTGCTGAGTGAACAAACCTCCAGAGACTGTG 921
DB 276 ThrThrThrSerAlAlAlAlMetValHisGlyLeuValGlnLysThrGlnGlnLysLeu 295
QY 922 ACCCTGAGCCAGAAACAAAGCTTTTCACT----- 951
DB 296 GluThrThrLeuLysMetLysPheLeuSerLeuAspLeuGlnGlyAspProArgLysSer 315
QY 952 -----CGAGTGTGATGAGTCTTTCTTCTTAATGAGCCTGTGATCTA 996
DB 316 GlnProLeuLeuGlyThrValIleAspGlyMetLeuLeuLeuLysThrProGlnGlnLeu 335
QY 997 TTGTCTCAAGAAAGCATTTAAAGCAATTCCTTCATCATCGGATCAATTAACAGAGTGT 1056
DB 336 GlnAlAGlnArgAsnPheHisThrValProLysMetValGlyIleAsnLysGlnGlnPhe 355
QY 1057 GCGTTCGCTGCTGCTATGAAGAGAGCTCTCGAGATCTCGAGTGGCTCC-----AACAAG 1110
DB 356 GlyTrpLeuLysLeuProMetGlnLeuMetSerLysProLeuValCysIleAlAlLysGlnLeu 375
QY 1111 TCCCTTCCCTTCATCTGATACAAAC-----ATCCGACATCCGCGCTCAAGT 1161
DB 376 LysThrAlAlMetSerLeuLeuLeuTrpLysSerLysProLeuValCysIleAlAlLysGlnLeu 395
QY 1162 TTGACCTTGGCTGATGAATTAATCTTCATGACAGACACTCCCTGACTGAATAATCCAGAC 1221
DB 396 IleProGlnAlAlThrGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 415
QY 1222 AGTCTTGTGACCTGTGAGATGTGTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1281
DB 416 LeuPheLeuAspLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 435
QY 1282 TATACACAGATGTGTGTGCACTGTCTACTTCTATGAGTTTGGCAGCGGCTCAAGTGC 1341
DB 436 AsnHisArgAspAlAGlyAlAlProThrLysMetLysGlnPheGlnLysArgProSerPhe 455
QY 1342 TTGAAGACAGAAAGCGGCTTTTTCGAAGCCGACCGACCTGATGAAGTCCGCTTGTG 1401
DB 456 SerSerAspMetLysProLysThrValIleGlyAspHisGlyAspGlnLeuPheSerVal 475
QY 1402 TTGCTGTGCTCTTCTGAAAGGGGACATTTATGTTTCGAAGAGCCAGAGAGAGAG 1461
DB 476 PheGlyAlAlProPheLeuLys-----GlnGlyLysSerGlnGlnGln 489
QY 1462 AAGTTACTGACCGGAAAGATGAATAATCTGGGCTACCTTTGCTGGAACCGGGAATCCT 1521
DB 490 IleArgLysSerLysMetValMetLysPheTrpAlAlAsnPheAlAlArgAsnGlyAsnPro 509
QY 1522 AATGGAACAGACCTGTCTGTGCGACGCTTATATCTGACTGAGAGAGTATCCCAAGCTG 1581
DB 510 AsnGlyGlnGlyLeuProHisLysTrpProGlnLysGlnLysGlnLysGlnLysGlnLys 529
QY 1582 GACTTGAACATGAGCCTCGGACAGAGACTCAAGAACCGGCGGTGATTTTGGACCAAG 1641
DB 530 GlyAlAlAsnThrGlnAlAlAlAGlnLysLeuLysAspLysGlnValAlAlPheTrpThrAsn 549
```

RESULT 10
S62788
carboxylesterase (EC 3.1.1.1) ES-4 precursor, liver - rat
N:Alternate names: hydrolyase B
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Oct-1996 #sequence revision 13-Mar-1997 #ext_change 09-Jul-2004
C:Accession: S62788; S51203; A55304; S49257
R:Robb, M.; van Schaftingen, E.; Beaufay, H.
Biochem. J. 313, 821-826, 1996

A:Title: Cloning and sequencing of rat liver carboxylesterase ES-4 (microsomal palmitoyl
A:Reference number: S62788; MUID:96190723; PMID:8611161
A:Accession: S62788
A:Molecule type: mRNA
A:Residues: 1-561 <RO>
A:Cross-references: UNIPROT:064573; EMBL:X81825; NID:9550417; PDB:CAA57419.1; PID:95504
A:Experimental source: liver
A:Organism: Rattus norvegicus
A:Arch. Biochem. Biophys. 315, 495-512, 1994
A:Title: Purification and characterization of two rat liver microsomal carboxylesterases
A:Reference number: S51202; MUID:95077430; PMID:7986098
A:Accession: S51203
A:Molecule type: protein
A:Residues: 19-48 <MO>
A:Experimental source: liver
A:Organism: Rattus norvegicus
A:Arch. Biochem. Biophys. 269, 23688-23696, 1994
A:Title: Rat kidney carboxylesterase. Cloning, sequencing, cellular localization, and re
A:Reference number: A55304; MUID:95050819; PMID:7961958
A:Accession: A55304
A:Molecule type: mRNA
A:Residues: 1-6, 'P', 8-59, 'P', 61-212, 'A', 214-252, 'T', 254-309, 'IT', 312-341, 'N', 343-424, 'P'
A:Cross-references: GS:U010697; NID:9562007
A:Note: the sequence in Genbank entry RNU010697, release 107, (PID:9562008) has the codon
R,Robbi, M.; Beaufay, H.
submitted to the EMBL Data Library, September 1994
A:Reference number: S49257
A:Accession: S49257
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-561 <RO>
A:Cross-references: EMBL:X81825; NID:9550417; PDB:CAA57419.1; PID:9550418
A:Function: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
A:Superfamily: cholinesterase; cholinesterase homology
A:Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
P.118/Domain: signal sequence #status predicted <SIG>
P.119-561/Product: carboxylesterase ES-4 #status experimental <MAT>
P.150-551/Domain: cholinesterase homology <CHE>
P.1221,466/Active site: Ser, His #status predicted
P.301/Binding site: carboxylate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.:	4,62e-80	Length:	561
Score:	1139.00	Matches:	243
Percent Similarity:	58.04%	Conservative:	82
Best Local Similarity:	43.39%	Mismatches:	195
Query Match:	35.93%	Indels:	40
DB:	2	Gaps:	10

US-10-023-515-3 (1-1746) x S62788 (1-561)

```
QY 43 TTCTGATTTCTCCAGCCCTTTGGAGACAGACAGTGGGGAATACTGGCCTTCTGCT 102
DB 5 PheLeuIleLeuValSerLeuAlaThrCysValValTyrGly-----AsnProSerSer 22
QY 103 GAAAGGCCACAGAGACACAGCGCTGGAGTCACTGAGGACAGACAGTCACTGCTG 162
DB 23 Pro--ProValValAspThrThrIlysgIlyValLeuGlyIlyrValSerLeuGlu 41
QY 163 GGAAGCCCTGGCTGCTGAGAGTTCCTCGGAGCCCTTTGCTGCTCCCGCGCTGGGA 222
DB 42 GAlValThrGlnSerValAlaValAlaPheLeuGlyValProPheAlaIlyProLeuGly 61
QY 223 TCCCTGACATTTACGACCCGAGCGCTGATCGCCCTGGATATCTGGACAGAACACC 282
DB 62 SerLeuAlaGlnPheAlaProProGlnProAlaGluProIlyPheSerPheValIlyAsnThr 81
QY 283 TCTTACCTCAATTTGCTGCTCAGAACTCAGAGTG-----CTGCTCTTA 327
DB 82 ThrTyrProProIlyMetCysSerGlnAspAlaAlaIlysgIlyGlnArgMetAsnAspLeu 101
QY 328 GATCAACAGATGCTCAAGTGTGATTAACCGAAATTCGAGGTGTCAAGAGACGCTCTAC 387
```

```
DB 102 ThrAsnArgIlyGlyIlyIleHis-----LeuGlnPheSerGlyAspCysLeuTyr 118
QY 388 CTGAACATCTTTCGCCCTCCACCCGATACAGGCTCAAGCCTCCGCTTGGATGG 447
DB 119 LeuAlaIlyTyrThrProAlaAspPheThrIlyAsnSerArgLeuProValMetValTyr 138
QY 448 TTCCAGAGAGGCTTCAAGACTGCTCAGCGCTCACTTGTGATGGGTCCGCTGGCT 507
DB 139 IleHisGlyGlyGlyMetThrLeuGlyIlyAlaSerThrTyrAspGlyArgValLeuSer 158
QY 508 GCTTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
DB 159 AlaTyrGlnAsnValValValAlaIleGlnTyrArgLeuGlyIlyIleProIlyPhePhe 178
QY 568 ACCAATGAGATGACAGATGCTCCGCGGAGATCGGCTTCAAGAGACAGTGGCTGCTG 627
DB 179 SerThrIlyAspGlyHisSerArgIlyAsnThrGlyHisIleAspGlnValAlaIleLeu 198
QY 628 TCCCTGGCTCAGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 687
DB 199 HisThrValGlnAspAsnIleAlaAsnPheGlyIlyAspProGlySerValThrIlePhe 218
QY 688 GCGAGTCCGCGGAGCCATTAAGTGTCTTAAGTCTTAAGTCTTCCAGTGGCCAAAGC 747
DB 219 GlyIlySerIlyGlyIlyPheSerValSerValIleValIleLeuSerProLeuThrIlyAsn 238
QY 748 TTATTCACAAAGCATATGAGAGTGGGTGGGATCATCTTCACTGAGGCCAT 807
DB 239 LeuPheHisArgAlaIleSerIlySerIlyValValIlePheLeuProGlyLeuLeuThrIly 258
QY 808 GATTATGAGAGAGAGAGACCTCGAGGTGGTGGACATTTCTGTTGTAACAATGCTGA 867
DB 259 AspValAlaGProAla-----AlaIlyGlnIleAlaAspMetAlaGlyCysGlyThrThr 276
QY 868 GACTGTGAGGCTTCTGAGTGTGCTGAGACAAACCTCCAGAGAGCTGTGACCTG 927
DB 277 ThrSerAlaIleIleValHisIleCysLeuArgIlyIlyThrGlnGlnGlnIleLeuGlnIle 296
QY 928 AGCCGAAACAAAG----- 942
DB 297 MetIlyValMetAsnLeuIleIlyLeuSerSerGlnArgAspAsnIlyGlySerTyrHis 316
QY 943 TCTTTCATCGAGTGGTATGATGCTTCTTCTTAATGAGCCTCTGATATATGCT 1002
DB 317 PheLeuSerThrValValAspAsnValIleLeuProIlyAspProIlyGlyIlyIleValAla 336
QY 1003 CAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATTAACAGAGTGGCTTC 1062
DB 337 GlnIlyAsnPheAsnThrValProTyrIleValGlyIleAsnIlyGlnIlyGlySerIlyTyr 356
QY 1063 CTGCTGCTT-----ATGAGAGGCTCTGATGATCTCAAGTGGCTCAAGATCCCTT 1116
DB 357 LeuLeuProIlyMetMetGlyPheValIleProAlaAspValGlu---LeuAspIlySerMet 375
QY 1117 GCCCTTCATCTGATCA-----AACATCTCGACATCCGCGCTCAAGATTTGAC 1167
DB 376 AlaIleThrIleLeuGlnIlyPheAlaSerLeuTyrGlyIleProGlnAspIleIlePro 395
QY 1168 CTGTGGCTAATGAATCTTCATGACAGACCTCCCTGATCAATTCGAGACAGTCTT 1227
DB 396 ValAlaIleGlyIlyTyrArgIlySerArgAspAsnIleIlyIleArgAspIlyIle 415
QY 1228 CTGACCTGCTGAGATGCTTCTTGTGCTCTGATCATGATCATGACAGCTGATATC 1287
DB 416 LeuAlaPheIleGlyAspValSerPheSerIleProSerValMetValSerArgAspHis 435
QY 1288 AGAGATGCTGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1347
DB 436 ArgAspAlaGlyAlaProThrTyrMetCylArgIlyTyrIlyTyrProSerPheSerSer 455
QY 1348 GACACGAGCCGCTTTGTCAAGAGCCGACACGCTGATGAGTCCGCTTGTGTGCT 1407
```

Db 456 ProGlnArgProLysHisValValGlyAerPHiValAerAspLeuTyrSerValPheGly 475
QY 1408 GGTGCTTCTGTAAGGGGACATTTGTATGTTCCAGAGGCCACGGAGAGAGAACTTA 1467
Db 476 AlaProIleLeuArg-----AerGlyAlaSerGlnGlnIleVal 489
QY 1468 CTGACCCGGAAGATATGAAATACTAGGAGTCACTTTGCTGCAACCGGAACTTAATGGG 1527
Db 490 LeuSerIlyMetValMetLysPheThrAlaAenPheAlaArgAenGlyAsnProAenGly 509
QY 1528 AACGACCTGTCTGTGGCCAGCTTAATATCTGACTGAGACGATCTCCAGCTGACTTG 1587
Db 510 ArgGlyLeuProHisIleTyrProGlnTyrAerGlnTyrGlnGlnTyrLeuGlnIleGlyAla 529
QY 1588 AACATGAGCCTCGACAGAGACTCAAAAGAACCGGGGTGGATTTTGGACCAAGCAACTC 1647
Db 530 ThrThrGlnInSerGlnArgLeuLysValGlnGlnValAlaPheTyrThrGlnLeuLeu 549
RESULT 11
A55281
carboxylesterase (BC 3.1.1.1) egasyn - mouse
N:Alternate names: beta-glucuronidase endoplasmic reticulum-targeting protein; esterase-
C:Species: Mus musculus (house mouse)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004
C:Accession: A55281
R:Ovnic, M.; Swanik, R.T.; Fletcher, C.; Zhen, L.; Novak, E.K.; Baumann, H.; Heintz, N.;
Genomics 11, 956-967, 1991
A:Title: Characterization and functional expression of a cDNA encoding egasyn (esterase-
A:Reference number: A55281; MUID:92147141; PMID:1783403
A:Accession: A55281
A:Status: preliminary
A:Molecule type: mRNA, protein
A:Residues: 1-562 <OVN>
A:Cross-references: UNIPROT:064176; GB:S80191; NID:9244727; PIDN:AA821335.1; PID:9244728
A>Note: sequence extracted from NCBI backbone (NCBI:80191, NCBI:80194)
C:Superfamily: cholinesterase; cholinesterase homology
C:Keywords: carboxylic ester hydrolase; endoplasmic reticulum
F:51-552/Domain: cholinesterase homology <CHS>
F:222,467/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 5,06e-80 Length: 562
Score: 1138.50 Matches: 240
Percent Similarity: 58.52% Conservative: 76
Best Local Similarity: 44.44% Mismatches: 195
Query Match: 35.91% Indels: 29
DB: Gaps: 6

US-10-023-515-3 (1-1746) x A55281 (1-562)

QY 91 GGGCTTCTGCTGAAGGGCCACAGAGAACCAAGCTGGATGGATTCAAGGCAAGCA 150
Db 19 GLyHisProSerSerProProMetValAspThrValGlnGlyLysValLeuGlyLysTyr 38
QY 151 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTGTTCTCGAGTCCCTTGTGCT 210
Db 39 lIeSerLeuGlnGlyPheThrGlnProValAlaValPheLeuGlyValProPheAlaLys 58
QY 211 CCCCCGTGGATCCCTCGGATTTTACGAAACCCGAGCTGATGCGCTCGGAAATAC 270
Db 59 ProProLeuGlnGlySerLeuArgPheAlaProProGlnProAlaGlnProTyrPheSerVal 78
QY 271 CGAAGAACCACTCTCACTCCCTAATTTGCTGCTCCAGAACTCAGATGGCTCTCTAGAT 330
Db 79 LysAsnAlaThrSerTyrProProMetCysPheGlnAspProValThrGlnIleVal 98
QY 331 CAACACAGCTCAAGGTGATTAACCGAAATTCGAGT-----TCAGAAAGTGCCTC 384
Db 99 AsnAspLeuLeuThrAsnArgLysGlyLysIleProLeuGlnPheSerGlnAspCysLeu 118
QY 385 TACCTGAACATTTATGGGCTGCCACGCGCATACAGGCTCCAGCTCCCGTCTTGCTG 444
Db 119 TyrLeuAsnIleTyrThrProAlaAspLeuThrLysSerAspArgLeuProValMetVal 138

QY 445 TGGTCCCAAGAGGTCCTTCAAGACTGCTCAGCTTCATCTTGTGATGGGTCGGCCTG 504
Db 139 TrpIleHisGlyGlyGlyLeuValLeuGlyValAlaSerThrTyrAerGlyLeuValLeu 158
QY 505 GCTGCTATGAGAGAGTGTGTTGGTGGCTCCAGTCCCGGCTTAGAAATATTTGGTTTC 564
Db 159 SerThrHisGlnAenValValValValValIleGlnTyrAerGlnGlyIleTyrGlyPhe 178
QY 565 TTCACCAATGGATCAGCATGCTCCGGGAATCTGGGCTTCAAGACCAAGATGGCTGCT 624
Db 179 PheSerThrGlyAerGlnHisSerArgLysAenThrGlnHisLysAerGlnValAlaAla 198
QY 625 CTGTCCGGGGCCAGAAAGATATGATGAGTCTTCGCTGGGACCCCACTCTGTACCATC 684
Db 199 LeuHisThrValGlnAerAsnIleAlaLysPheGlyGlyAerProGlySerValThrIle 218
QY 685 TTTGGAGAGTCCGGGAGCCATTAAGTGTCTTCTAGTCTTATATGCTGTCCATGGCCAA 744
Db 219 PheGlyGlnSerAlaGlyGlyGlnSerValSerValLeuValLeuSerProLeuAlaLys 238
QY 745 GGTATTTCCACAAAGCCATCATGAGAGATGGGGTGGCCATCATCTTACTTGAAGGCC 804
Db 239 AsnLeuPheGlnArgAlaIleSerGlnSerGlyAlaAlaLeuThrAlaGlyLeuValLys 258
QY 805 CATGATTATGAGAGATGAGAGACCTGACAGTGGTGTGCACATTTCTGTGTAACAATGCG 864
Db 259 LysAenThrArgProLeuAlaGlnLysIleAlaValIleSerGlyCysLysAen----- 276
QY 865 TCAGACTCTGAGGCTCTGAGTGGCTGAGAGCAAAACCTTCCAAAGAGCTGTGACC 924
Db 277 ThrThrSerAlaAlaMetValHisCysLeuArgGlnTyrThrGlnGlnGlnLeuGly 296
QY 925 CTCAGCCCAAAACAAAGTCTTTC----- 948
Db 297 ThrThrLeuLysLysLeuAenLeuPheLysLeuAerLeuHisGlyLysSerArgGlnSerHis 316
QY 949 -----ACTGAGGGTGTGATGATGTCCTTCTTCTTATAGAGCTTCATAGATCTATTG 999
Db 317 ProPheValProThrValLeuAerPheGlyValLeuLeuProLysMetProGlnGlnIleLeu 336
QY 1000 TCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCCGAGTCAATTAACAGAGTGTGCG 1059
Db 337 AlaglyLysAsnPheAenThrValProTyrIleValGlyLysAenLysGlnGlnPheGly 356
QY 1060 TTCTGTGCTGCT--ATGAAGAGGCTCTGAGATCTTCAGTGGCTCAACAAAGTCCCTT 1116
Db 357 TrpIleLeuProThrMetMetAsnTyrProProSerAspValLysLeuAerGlnMetThr 376
QY 1117 GCCCTCATCTGATACAAAC-----ATCCTGACATCCCGGCTCAGTATTTTGAC 1167
Db 377 AlaMetSerLeuLeuLysLysSerSerPheLeuLeuAenLeuProGlnAerAlaIleAla 396
QY 1168 CTGTGGCTAATGATACTTCATGACAGACACTCCCTGACTGAATTCGAGACAGTCTT 1227
Db 397 ValAlaIleGlyLysTyrLysLeuArgAspLysAspTyrThrGlyArgAenLysAerGlnLeu 416
QY 1228 CTGACCTCTTGGAGATGTGTTCTTGTGTGCTCCCTGCACTGATCACAGCTCGATATCAC 1287
Db 417 LeuGlnLeuIleGlyAspValAlaPheGlyValProSerValIleValSerArgGlyHis 436
QY 1288 AGAGATGCTGATGACCTGTCTACTTCTATGAGTTTCGAGACCGGCTCAGTCTTGAA 1347
Db 437 ArgAspAlaGlyAlaProThrTyrMetTyrGlnPheGlnTyrSerProSerPheSerSer 456
QY 1348 GACACGAAGCCGGCTTTGTCAAAGCCGACACAGCTGATGAATCCGCTTGTTCGGT 1407
Db 457 GluMetLysProAspThrValValGlyAerPHiGlyAspGlnIleTyrSerValPheGly 476
QY 1408 GTGCTTCTGTAAGGGGACATTTTATGTTTGAAGAGACCCAGAGAGAGAAAGTTA 1467
Db 477 AlaProIleLeuArgLys-----GlyThrSerGlnGlnIleAsn 490

Db 481 ArgAsnGlyAsnProAsnGlyGluGlyLeuProHisIleTrpProIleThrAspGlnIleVal 500
QY 1567 CAGTACCTTCAGCTGAGCTTGAACATGAGCTTCGAGACAGACTCAAGAACCAGCGGTG 1626
Db 501 GlyTyrLeuGlnIleGlyAlaThrThrGlnGlnAlaGlnIleValLeuGlyGlyGluVal 520
QY 1627 GATTTTGACAGCAGCACCATC 1647
Db 521 AlaPheTrpThrGluLeuLeu 527

RESULT 13
S71597
carboxylesterase (EC 3.1.1.1) precursor, liver - rat
N.Alternate names: hydrolase C
C.Species: Rattus norvegicus (Norway rat)
C.Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 13-Sep-1998
C.Accession: S71597
R.Yan, B.; Yang, D.; Parkinson, A.
Arch. Biochem. Biophys. 317, 222-234, 1995
A.Title: Cloning and expression of hydrolase C, a member of the rat carboxylesterase fam
A.Reference number: S71597; MUID:95177656; PMID:7872788
A.Accession: S71597
A.Status: not compared with conceptual translation
A.Molecule type: mRNA
A.Residues: 1-561 <YAN>
A.Experimental source: liver; endoplasmic reticulum
C.Function:
A.Description: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
C.Superfamily: cholinesterase; cholinesterase homology
C.Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-561/Product: carboxylesterase #status predicted <CAT>
F:50-551/Domain: cholinesterase homology <CHS>
F:558-561/Region: endoplasmic reticulum retention signal
F:779/301/Binding site: carbonylate (Asn) (covalent) #status predicted
F:721,466/Active site: Ser, His #status predicted

Alignment Scores:
Pred. No.: 2,526-79 Length: 561
Score: 1129.50 Matches: 236
Percent Similarity: 58.64% Conservative: 83
Best Local Similarity: 43.38% Mismatches: 188
Query Match: 35.63% Indels: 37
Gaps: 9

US-10-023-515-3 (1-1746) x S71597 (1-561)
QY 91 GGGCCCTTCTGCTGAAGGCCCAAGAGAACACCGCTGGATGATTCAAGGCAAGCA 150
Db 18 GlysAnPProSerSerProProValValAlaPheThrMetLysGlyLysValLeuGlyLysTyr 37
QY 151 GTCACTGCGCTGGAGAGCCCTGCTGTCGATGACGCTTCCCTCGAGTCCCTTGTGCTCT 210
Db 38 AlaSerLeuGlnGlyValThrGlnSerValAlaValPheLeuGlyValProPheAlaLys 57
QY 211 CCCCCGTGGATCCTGCGATTATTCGAACCCGAGCCTGCATCGCCCTGGATTAATCTG 270
Db 58 ProProLeuGlnGlySerLeuAlaPheAlaProProGlnProAlaGlnLysProTrpSerPheVal 77
QY 271 CGAGAACCACTCTCAATCTTAATTGCTGCTCGAAGCTCAAGAGTGG----- 318
Db 78 LysAnThrThrThrThrTyrProProMetCysSerGlnAspAlaThrLysGlyLysArgMet 97
QY 319 ---CTGCTTAGATCAACACATGCTCAAGTGAATTAACCGAAATTCGAGTCTCAAA 375
Db 98 AsnAspLeuLeuThrAsnAlaGlyLysGlnLysValHis-----LeuGlnPheSerGlu 114
QY 376 GACTGCTCTACTGAACATCTATGCGCTGCGCCAGCCGATACAGGCTCCAGCTCCC 435
Db 115 AspPheLeuTyrLeuAlaIleTyrThrProAlaAspPheThrLysAspSerArgMetPro 134
QY 436 GTCTTGATGTGTTCCAGAGAGTGTCTTCAAGACTGGCTCAGCTTCATCTTGATGG 495

Db 135 ValMetValTrpIleIleGlyGlyGlyLeuThrGlnGlnGlyAlaSerThrTyrAspGly 154
QY 496 TCCGCCCTGCTGCTATGAGACGTGCTGTGTGTGCTGTCAGATACCGCTAGAA 555
Db 155 ArgValLeuSerAlaTyrGlyAsnValValAlaValAlaIleGlnTyrArgLeuGlyLe 174
QY 556 TTGGTTCTTCACCAATGGGATCAGATGCTCCGGGAATCTGGCTTCAAGACAG 615
Db 175 TrpGlyPhePheSerThrGlyAspGlnHisSerArgLysAsnTrpGlyLysLeuAspGln 194
QY 616 GTGGCTGCTGTGCTGGATCCAGAAACATCAAGTTCCTGGTGGGACCCAGCTCT 675
Db 195 ValAlaAlaLeuHisTrpValGlnAspAsnIleAlaAsnPheGlyGlyAspProGlySer 214
QY 676 GTGACCATCTTGGCGAGTCCGCGAGCCATTAAGTGTTCCTAATCTTATCTGTCC 735
Db 215 ValThrIlePheGlyGlnSerAlaGlyLysPheSerValSerValLeuValLeuSerPro 234
QY 736 ATGGCCAAAGCTTATTCACAAAGCATCATGAGAGTGGGGTGGCCATATCCCTTAC 795
Db 235 LeuSerLysAsnLeuTyrThrAlaAlaIleSerGlnSerGlyValValLeuIleThrGln 254
QY 796 CTGAGAGCCCATGATTATGAGAAAGTAGAGACCTGAGGTGTGACATTTCTGTGT 855
Db 255 LeuPheThrLysAspValArgProAla-----AlaLysGlnIleAlaAspMetAlaGly 272
QY 856 AACATGCGTCAGACTTGAGCCCTGCTGAGTGGCTGAGACAAACCTCCAGAG 915
Db 273 CysLysThrThrThrSerAlaIleIleValHisCysLeuArgGlnLysThrGlnGln 292
QY 916 CTGTGACCCCTGACAGAA-----ACAAAG 942
Db 293 LeuLeuGlnIleMetCysGlyLysMetAsnLeuIleLysLeuSerGlnAlaAspThrLys 312
QY 943 TCT-----TTCACTCGAGTGTGATGTGCTTCTTCTTAATGAGCCCTTA 990
Db 313 GluSerTyrHisPheLeuSerThrValIleAspAspValValLeuProLysAspProLys 332
QY 991 GATCTATTTGCTCAGAAAGCATTTAAAGCATTCCTTCCATCATCGAGTCAATACAC 1050
Db 333 GlnIleLeuAlaGlnLysAsnPheAsnThrValProTyrIleValGlyIleAsnLysGln 352
QY 1051 GAGTGTGCTTCCCTGCTGCT-----ATGAGAGAGCTCTGAGATCCAGTGTCC 1104
Db 353 GluCysGlyTyrTrpLeuLeuProThrMetMetLysPheValProProAspValLys--Leu 371
QY 1105 AACAGTCCCTTGCCCTCATCTGATACAA-----AACATCTGCATCCGCT 1155
Db 372 AspLysLysMetAlaIleMetLeuLeuGlnLysPheAlaSerIleTyrGlyIleProGlu 391
QY 1156 CAGTATTGACCTTGCGCTAATGAAATCTTCCATGACAAAGACTCCCTGATGAATC 1215
Db 392 AspIleIleProValAlaIleGlnLysTyrArgLysGlySerAspAspProIleLysIle 411
QY 1216 CGAGACAGTCTTCTGAGCTTGGAGATGCTTCTTGTGCTCCCTGACATGATACA 1275
Db 412 ArgAspLysIleLeuAlaPheIleGlyAspValLeuPheCysIleProSerValMetVal 491
QY 1276 GCTCGATATCAAGAACTGCTGCACTGTCTACTTATGAGTTTGGCACCGGCT 1335
Db 432 SerArgAspHisArgAspAlaGlyAlaProThrTyrValTyrGlnTyrGlnTyrPro 451
QY 1336 CAGTGTTCGAAGACAAAGACCGGCTTTGTCAAAAGCCAGACGCTGATGAAGTCCG 1395
Db 452 SerPheSerSerProGlnArgProLysAspValValGlyAspHisAlaAspAspValTyr 471
QY 1396 TTGTGTTCGTGGTGTCTTGAAGGGGACATTTGATGTCAGAGCCAGGAG 1455
Db 472 SerValPheLysLysAlaProIleLeuArg-----AspGlyAlaSerGlu 485
QY 1456 GAGAGAAATTACTGAGCCGGAAGATGATGAATACTAGGCTACTTGTCTCGAACCGG 1515
Db 486 GlnGlnIleLysLeuSerLysMetValMetLysPheTrpAlaAsnPheAlaArgAsnGly 505

QY 1516 AATCTAATGGAACGACCTGCTCTGTCGACGCTTAATCTGACTGACGACCTC 1575
 |||||
 Db 506 AsnProhema1AaAggLyLeuProHisrPrroGlnrYrAaBgInLygluInLyriLeu 525
 |||||
 QY 1576 CAGCTGACCTGAACATGAGCCCTCGACAGACACTCAAGAACCGCGGCTGATTTTGG 1635
 |||||
 Db 526 GlnIleGlyAlaThrThrGlnGlnSerGlnAArgLeuLySaIaGluGluValAlaIaPheTrp 545
 |||||
 QY 1636 ACCGACCAATC 1647
 |||||
 Db 546 ThrGlnLeuLeu 549
 |||||

RESULT 14

JC2447
 carboxylesterase (EC 3.1.1.1) ES-3 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 29-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
 C:Accession: JC2447; S23462
 R:Robbl, M.; Beaufay, H.
 Biochem. Biophys. Res. Commun. 203, 1404-1411, 1994
 A:Title: Cloning and sequencing of rat liver carboxylesterase ES-3 (EGASYN).
 A:Reference number: JC2447; MUID:95032008; PMID:7945287
 A:Accession: JC2447
 A:Molecule type: mRNA
 A:Residues: 1-561 <ROB>
 A:Cross-references: UNIPROT:Q63108; GB:X81395; NID:9550146; PIDN:CA57158.1; PID:9550147
 A:Experimental source: Liver
 R:Medda, S.; Proia, R.L.
 Eur. J. Biochem. 206, 801-806, 1992
 A:Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting h
 A:Reference number: S23460; MUID:9229908; PMID:1606962
 A:Accession: S23462
 A:Molecule type: translation not shown
 A:Status: preliminary
 A:Residues: 503-554, 'R', 556-561 <MED>
 A:Cross-references: EMBL:X65295; NID:957557; PIDN:CAA46390.1; PID:957558
 C:Superfamily: cholinesterase; cholinesterase homology
 C:Keywords: carboxylic ester hydrolase; glycoprotein
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-551/Product: carboxylesterase ES-3 #status predicted <MAT>
 F:79,107,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:221,466/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.: 8,04e-79 Length: 561
 Score: 1123.00 Matches: 245
 Percent Similarity: 57.07% Conservative: 86
 Best Local Similarity: 42.24% Mismatches: 193
 Query Match: 35.43% Indels: 56
 DB: 2 Gaps: 11

US-10-023-515-3 (1-1746) x JC2447 (1-561)

QY 37 TGCCTTTTCCGATTCCTCAGCCCTCTGTTGGACACAGACAGTGGGAAAAAATGGGCT 96
 |||||
 Db 2 CysLeuTrpAlaLeuIleLeuValPheLeuAlaIaPheThrAlaGly-----GlyHis 19
 |||||
 QY 97 TCTCTGAAGGGCCACAGAGAACAGAGGCTGGATGATTCAGGACGACGACTACT 156
 |||||
 Db 20 ProSerSerLeuProValAlaPheThrLeuGlnGlyValLeuGlyLysrYrValSer 39
 |||||
 QY 157 GTGCTGGAGACCTCTGCTCTGTGAAGCTGTTCTCGAGAGTCCCTTTGCTGCTCCCG 216
 |||||
 Db 40 LeuGluGlyPheThrGlnProValAlaValPheLeuGlyValProPheAlaIaPhePro 59
 |||||
 QY 217 CTGGATCCCTGGCATTTACAGACCGGAGCTGATGGCCCTGGGATTAATTGGAGAA 276
 |||||
 Db 60 LeuGlySerLeuAaPheAlaProProGlnProAlaGluProTrpSerPheValIaAsn 79
 |||||
 QY 277 GCACCTCTCACTCAATTTGTGCTCCAGAAC----- 309
 |||||

Db 80 ThrTrpSerTrpProPheMetCysSerGlnAspProValAlaGlyGlnIleValAsnAsp 99
 QY 310 -----TCAGAGTGGCTGCTCTTAGATCAACAGACTCAAGGTGCATTAACCGAAATTC 363
 |||||
 Db 100 LeuLeuTrpAsnTrp-----GluGluAsnIleSerLeuGlnPhe----- 112
 |||||
 QY 364 GAGAGTGCAGAAAGCTGCTCTTACCTGAACATCTATGGCCCTGCCACGCCGATACAGGC 423
 |||||
 Db 113 -----SerGluAspCysLeuTrpLeuAsnIleTrpProAlaAspLeuThrIaAsp 130
 |||||
 QY 424 TCAGAGCTCCCGCTTTGCTGCTGCTCCAGAGAGTGCCTTCAGACTGCTCAACCTCC 483
 |||||
 Db 131 AspArgLeuProValMetValTrpIleHisGlyGlyGlyLeuValLeuGlyValAsp 150
 |||||
 QY 484 ATCTTGATGGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 |||||
 Db 151 ThrTrpAspGlyLeuAlaLeuSerThrHisGluAsnValValValValIleGlnTrp 170
 |||||
 QY 544 CGGCTAGGAATATTTGTTGTTTCTTCAACACATGGATCAGACTGCTCCGCGGAATCGGCC 603
 |||||
 Db 171 ArgLeuGlyIleTrpGlyPhePheSerThrGlyAspGlnHisSerArgGlyAsnTrpGly 190
 |||||
 QY 604 TTCAGAGACCGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
 |||||
 Db 191 HisLeuAspGlnValAlaAlaLeuHisTrpValGlnAspAsnIleAspAspPheGly 210
 |||||
 QY 664 GACCCACCTCTGATCCATCTGTTGGAGTCCGCGGAGGCATATGATGTTCTGCTGCT 723
 |||||
 Db 211 AspProGlySerValThrIlePheHisGlyLeuSerAlaGlyGlyGlySerValSerVal 230
 |||||
 QY 724 ATACTGTCTCCATGAGCCCAAGAGCTTATTCACAAAGCATATGAGAGTGGGCTGAGCC 783
 |||||
 Db 231 ValLeuSerProLeuAlaIaValAsnLeuPheHisValIleSerGlyValAla 250
 |||||
 QY 784 ATCATCTCTTACCTG--GAGGCCATATATATGAAAGTGAAGTCAAGTGGCT 840
 |||||
 Db 251 LeuThrAlaGlyLeuValIaValLysAsnThrArgProLeuAlaGluValIleAlaVal 270
 |||||
 QY 841 GCACATTTCTGTGTAACAATGCGCTGAGACTGAGCCCTGAGGCTGAGGACA 900
 |||||
 Db 271 Ser-----GlyCysLysSerThrThrSerAlaSerMetValHisCysLeuArgGln 287
 |||||
 QY 901 AAACCTTCAGAGAGCTGCTGACCTTCAGCAGAAACAAAGTCTTCACT----- 951
 |||||
 Db 288 LysThrGluGluGluLeuLeuGlnThrThrLeuLysLeuAsnLeuPheSerLeuAspLeu 307
 |||||
 QY 952 -----GAGTGTGATGCTGCTTCTTT 975
 |||||
 Db 308 HisGlyAspSerArgGlnSerTrpProPheValProThrValLeuAspGlyValValLeu 327
 |||||
 QY 976 CCTAATGAGCTCTTATGATCTATTGCTCAGAAAGCATTAAGCAATTCCTTCATCATC 1035
 |||||
 Db 328 ProLysMetProGluGluIleLeuAlaGluLysAspPheAsnThrValProTrpTrpVal 347
 |||||
 QY 1036 GAGATCAATTAACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
 |||||
 Db 348 GlyIleAsnLysGlnGluPheGlyTrpIleLeuProThrMetLeuAsnTrpProProSer 367
 |||||
 QY 1093 CTCACTGCTCCACAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1143
 |||||
 Db 368 AspMetLysLeuAspProMetThrAlaThrSerLeuLysLysSerSerPheLeuLeu 387
 |||||
 QY 1144 CACATCCCGCTCAGATATTGACCTTGTGCTAATGAATCTTCATGACAGACAGACTCC 1203
 |||||
 Db 388 AsnLeuProGluGluAlaIleProValAlaValGluLysTrpLeuAspHisThrAspAsp 407
 |||||
 QY 1204 CTGACTGAATCCGAGAGTCTTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263
 |||||
 Db 408 ProAspArgAsnLysAspGlnLeuLeuGlnIleGlyAspValIlePheGlyValPro 427
 |||||
 QY 1264 GCATGATCAAGCTCGATATATCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1323
 |||||
 Db 428 SerValIleValSerArgGlyHisArgAspAlaGlyAlaArgThrTyMetTrpGluPhe 447
 |||||

```
OY 1324 CGGACCGGCTCAGTGTCTTGAGACAGAGCGGCTTTTGCAAGCCGACGCT 1383
      ::::::::::::::::::::|
Db 448 GlnTYrArgProSerheSerheSerlyMeTlyPProSerThValAlaGlymphISgly 467
OY 1384 GATGAAGTCCGCTTGTGTGTGCGTGTGCTTCCTGAGAGGGGACATTTGTTGCA 1443
      ::::::::::::::::::::|
Db 468 AaPgluIleTYrSerValPheGlyAlaProIleLeuAaGly----- 481
OY 1444 GGAGCCACGGAGAGAGAAATTACTGAGCCGGAAGATGATGAATACTGGCTACTTT 1503
      ::::::::::::::::::::|
Db 482 GlyThrSerLySerGlyGluIleValasnLeuSerLyMeMeMeTlyPheTrpAlaSnPhe 501
OY 1504 GCTCGAACCGGGGAATCCTTAATGGGAAAGACCTGTCTGTGGCGAGCTTAATGTA 1563
      ::::::::::::::::::::|
Db 502 AlaAaGaaNglyAasnProasnGlynglynglyLeuProhITrProGluTYrAaPglu 521
OY 1564 GAGCAGTACTTCAGCTGACTTGAACATGAGCTTCGACAGACACTCAAGACCGCG 1623
      ::::::::::::::::::::|
Db 522 GluGlyTYrLeuGlnIleGlyAlaThrThrgInGlnAlaGlnIleuGlySerGly 541
OY 1624 GTGAGTTTTTGACACAGACCATCCCGTGAATCTGTGCTCGACATGCTCCACAT 1683
      ::::::::::::::::::::|
Db 542 ValAlaPheTrpSerGlu-----LeuLeuAlaMeTlySerProLeuHISAla 556

RESULT 15
S10367
carboxylesterase (BC 3.1.1.1) ES-10 precursor, microsomal - rat
N.Alternate names: hydrolase A
C.Species: Rattus norvegicus (Norway rat)
C.Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text_change 09-Jul-2004
C.Accession: S10367; S12468; S51202; S23460; S14361
R.Robbl, M.; Beaufay, H.; Octave, J.N.
Biochem. J. 269, 451-458, 1990
A>Title: Nucleotide sequence of cDNA coding for rat liver PI 6.1 esterase (BS-10), a car
A.Reference number: S10367; MUID:90351366; PMID:2386485
A.Accession: S10367
A.Molecule type: mRNA
A.Residues: 1-565 <ROB1>
A/Cross-references: UNIPROT:O9R135; EMBL:X51974
A/Note: 168-Gln, 247-Tyr, 423-Met, and 506-Aasn were also found
R/Note: the sequence from Fig. 4 is inconsistent with that from Fig. 5 in having 265-Lys
R.Robbl, M.
submitted to the EMBL Data Library, February 1990
A/Reference number: S12468
A.Accession: S12468
A.Molecule type: mRNA
A.Residues: 1-264, 'K', 266-565 <ROB2>
A/Cross-references: EMBL:X51974; NID:956898; PIDN:CAA36236.1; PID:956899
R/Morgan, B.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
Arch. Biochem. Biophys. 315, 495-512, 1994
A>Title: Purification and characterization of two rat liver microsomal carboxylesterases
A/Reference number: S51202; MUID:95077430; PMID:7986098
A.Accession: S51202
A.Molecule type: Protein
A.Residues: 19-48 <MOR>
R/Medda, S.; Prota, R.L.
Eur. J. Biochem. 206, 801-806, 1992
A>Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting t
A/Reference number: S23460; MUID:92299008; PMID:1606962
A.Accession: S23460
A/Status: preliminary; translation not shown
A.Molecule type: mRNA
A/Residues: 1-185, 'Q', 187-422, 'M', 424-505, 'N', 507-565 <MED>
A/Cross-references: EMBL:X65296; NID:957553; PIDN:CAA46391.1; PID:957554
R/Gaustad, R.; Sleiten, K.; Lowhaug, D.; Fornum, F.
Biochem. J. 274, 693-697, 1991
A>Title: Purification and characterization of carboxylesterases from rat lung.
A/Reference number: S14361; MUID:91190080; PMID:2012599
A.Accession: S14361
A.Molecule type: protein
A/Residues: 19-26, 'D', 28-37 <GAU>
C/Superfamily: cholinesterase; cholinesterase homology
```

C.Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-565/Product: carboxylesterase #status predicted <MAT>
F:50-551/Domain: cholinesterase homology <CH>
F:79,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:221,466/Active site: Ser, His #status predicted

Alignment Scores:

Pred. No.:	1-05e-78	Length:	565
Score:	1121.50	Matches:	245
Percent Similarity:	56.06%	Conservative:	79
Best Local Similarity:	42.39%	Mismatches:	199
Query Match:	2	Indels:	55
DB:		Gaps:	11

US-10-023-515-3 (1-1746) x S10367 (1-565)

```
OY 34 TGTGCTTTTCTGATTTCTCAGCCCTGTGGGACACAGATGGGGAATACTGG 93
      |||
Db 8 TrpLeuPhe-----LeuAlaIaCySerThrAlaTrpGlyTYr----- 19
OY 94 CTTTGTGTAAGGCCACAGAGAACACAGGCTGGATGATTCAGGGCAAGTC 153
      |||
Db 20 ProSerSerPro---ProValAlaSerThrValLySerGlyValLeuGlyTYrVal 38
OY 154 ACTGAGTGGGAGAGCCCTGTGCTGTGAACGTTTCTCGGAGTCCCTTGTGCTGCC 213
      ::::|
Db 39 AsnIleuGluGlyPheAlaGlnProValAlaAlaPheLeuGlyTyrProPheAlaLySer 58
OY 214 CGGTGTGATCCCTGCGATTTTACGAACCCGAGCTGCATGCCCTGGAGATTAATCTGCA 273
      |||
Db 59 ProLeuGlySerLeuAaGpPheAlaProProGlnProAlaGluProTrpAaSnPheValLyS 78
OY 274 GAAGCCACTCTTACCTTAATTGTGCTCCAGAAC-----TCA 312
      |||
Db 79 AsnThrThrSerTYrProPrometCySerGlnAaPheAlaValGlyGlyGlnValLeuSer 98
OY 313 GAGTGGCTGCTTAAATGACACAGATGCTGAAGGTGATTAACCGAAATTCGAGTCA 372
      |||
Db 99 GluLeuPheThrAsnAaGlyGlyAsnIleProLeuGlnPhe-----Ser 113
OY 373 GAAGACTGCTTACTGATGAACATCTATGCGCTGCCAGCCGATACAGGCTCCAGACTC 432
      |||
Db 114 GluAaPcyLeuTYrLeuAaSnValTYrThrProAlaSerPheuThrLyAaSerAaGlyLeu 133
OY 433 CCGCTCTGTGTGTGTTCCAGAGAGTCCCTTCAAGACTGCGCTCCATCTTGAT 492
      |||
Db 134 ProValMetValTrpIleHisGlyGlyLeuValAlaGlyAlaSerThrTYrAaP 153
OY 493 GGGTCCGCGCTGGCTGTATGAGAGATGATGCTGTGTCGTCGTCGCTCAAGCTCAAG 552
      |||
Db 154 GlyGlnValLeuSerAlaHisGlyAaSnValValValAlaThrIleGlnTYrAaGlyLeu 173
OY 553 ATATTGTGTTTTCACACATGGGATCAAGATGCTCCGGGGAATCTGGGCTTCAAGAC 612
      |||
Db 174 IleTrpGlyPhePheSerThrGlyAaPgluHISerAaGlyIleuTrpGlyHisAaAaP 193
OY 613 CAGTGTGCTGTCTGTCTTGGTCCAGAAAGATGAGTCTTGGTGGGAGCCCAAGC 672
      |||
Db 194 GlnAlaAlaIleuHisTrpValGlnAaPAsnIleAlaSerPheGlyAaPProGly 213
OY 673 TGTGTGACATCTTTGGCGAGTCCGCGGAGCCATGAATGTTCTTAATCAATGCTCT 732
      |||
Db 214 SerValAlaThrIlePheGlyGluSerAlaGlyGlyPheSerValSerAlaIleuValLeuSer 233
OY 733 CCATGGCCAAAGGCTTATTCACAAAGCCATGAGAGAGTGGGCTGGCCATCATCCCT 792
      |||
Db 234 ProLeuAlaLyAaSnLeuPheHisAaGAlaIleSerGluSerGlyValValLeuThrSer 253
OY 793 TACTGAGGCCCATATTAATGAAAGAGAGAGACTGAGGTGTTGACATTTCTGT 852
      |||
Db 254 AlaLeuIleThrThrAaPserLySProIleAlaAaSnLeu-----IleAlaThrLeuSer 271
```

```
OY 853 GGTAAACAATGCGTCAGACTCTGAGCGCCCTGAGTGCCTGAGACAAACCTTCAAG 912
|||
|||
Db 272 GlyCysIsthrThrSerIaIaMetValHisCysLeuArgGlnIsthrGluAsp 291
|||
|||
OY 913 GAGCTGTGACCTCAGCCAGAAAACAAGTCTTTC----- 948
|||
|||
Db 292 GluLeuLeuGluThrSerLeuIsthrLeuAsnLeuPheIsthrLeuAspLeuGluIsthrPro 311
|||
|||
OY 949 -----ACTGAGTGTGATGGCTTCTTCTTCAATGAGCCT 987
|||
|||
Db 312 LysGluSerIsthrProPheLeuProThrValIleAspGlyValIleuProIsthrPro 331
|||
|||
OY 988 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATCGAGTCAATAAC 1047
|||||
|||||
Db 332 GluGluIleLeuAlaGluIsthrSerPheAsnThrValProIsthrIleValGlyIleAsnIsthr 351
|||||
|||||
OY 1048 CACGAGTGTGCTTCCTGCTGCTATGAAAGAGGCTCCTGAGATCCTCAGTGGC----- 1101
|||
|||
Db 352 GlnGluPheGlyIsthrIleProThrIleuMetGlyIsthrProIsthrSerGlnGlyIsthrLeu 371
|||
|||
OY 1102 -----TCCAACAAGTCCCTTGCCCTCATGTATACAAAACATCCTGCACATCCCG 1152
|||
|||
Db 372 AspGlnIsthrIsthrAlaIsthrSerLeu---LeuIsthrIsthrSerIsthrProThrIsthrIsthrSer 390
|||
|||
OY 1153 CCTCAGTATTGTGACCTTGTGCTATGATTAATCTTCATGACAAAGCATCCTGACTGAA 1212
|||
|||
Db 391 GluIsthrMetIleProValIsthrAlaGluIsthrIsthrPheGlyIsthrAspAspProAlaIsthr 410
|||
|||
OY 1213 ATCCGAGACAGTCTTCTGAGCTTGTGCTGAGATGTGTTCTTGTGGTCCCTGCACTGATC 1272
|||
|||
Db 411 ArgIsthrAspLeuPheGlnAspLeuValIleAspValIlePheGlyValIsthrSerValMet 430
|||
|||
OY 1273 ACAGCTCGATATCAGAGATGCTGTGCACTGTCTACTTCTATGAGATTTCGGCACCGG 1332
|||
|||
Db 431 ValSerArgSerHisArgAspAlaGlyIsthrProThrPheMetIsthrGluIsthrArg 450
|||
|||
OY 1333 CCTCAGTCTTGAAGACGCGGCTTGTCAAGCCGACGCGCTGATGAGTGC 1392
|||
|||
Db 451 ProSerPheValSerAlaMetArgProIsthrValIleGlyAspHisGlyAspGluLeu 470
|||
|||
OY 1393 CGCTTGTGTTGCGGTGCGCTTCTGTAAGGGGGAATTGTTATGTTCGAAAGAGCCACG 1452
|||
|||
Db 471 PheSerValPheGlySerProPheLeuIsthr-----AspGlyAlaSer 484
|||
|||
OY 1453 GAGAGAGAAAGTACTGAGCCGGAAGATGAATAATCTGGGCTACTTGTCTGAGACC 1512
|||
|||
Db 485 GluGluGluIsthrAsnLeuSerIsthrMetValMetIsthrIsthrIsthrIsthrIsthrIsthr 504
|||
|||
OY 1513 GGGATCTTAATGGGAAGCACTGTCTCTGTGGCCAGCTTAATCTGACTGAGCAGTAC 1572
|||
|||
Db 505 GlySerProAsnGlyGlyLeuProHisIsthrProGluIsthrAspGlnIsthrGluIsthrIsthr 524
|||
|||
OY 1573 CTCACAGTGAATTAACATGAGCTCGGACAGAGACTCAAGAAACCGCGGTGATTTT 1632
|||
|||
Db 525 LeuIsthrIleGlyAlaSerIsthrGlnIsthrAlaGlnIsthrLeuIsthrAspIsthrValIsthrIsthr 544
|||
|||
OY 1633 TGGACCAAGACCATCCCTGATCTCTGTCTCCGACATGCTCCACAGTCTCT 1686
|||
|||
Db 545 TrpSerGlu-----LeuArgAlaIsthrGluAlaIsthrGluIsthrPro 557
|||
|||
```

Search completed: June 15, 2005, 08:55:41
Job time : 60.53 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 02:39:48 ; Search time 155.861 Seconds
(without alignments)
11472.925 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 3170
Sequence: 1 atgcacacagggagctactctc.....ttctcttctgtcctctga 1746

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 segs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=x1h
-Q=/cgn2.1/USPTO.spool/6664091/runcat.14062005.133459.9335/app_query.fasta.1.4238
-DB=UniProt_03 -QFMT=fastaan -SUPFMT=n2p.rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEADSIZE=500 -MITLEN=0 -MAYLEN=2000000000
-USER=6664091@cgn.1.389 @runcat.14062005.133459.9335 -NCPU=6 -ICPU=3
-NO WMAP -LANG=QUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WAPN TIMEOUT=30 -THEREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2915	92.0	575	2	Q6NT32
2	2606	82.2	525	2	Q96DN9
3	2454	77.4	469	2	Q8N8C8
4	2250	71.0	575	2	Q6AW47
5	2094.5	66.1	575	2	Q6AW46
6	1985	62.6	545	2	Q81034
7	1676.5	52.9	361	2	Q95KH3
8	1298.5	41.0	557	1	SASB ANAPL
9	1282.5	40.5	561	2	Q91WGO
10	1281	40.4	559	2	Q8BK48
11	1276.5	40.3	559	2	Q35533
12	1271	40.1	559	2	Q8BM97
13	1266.5	39.9	556	2	Q6PDB7
14	1263.5	39.9	561	2	Q6GM54
15	1262.5	39.8	561	2	Q70177
16	1260.5	39.8	560	2	Q8R097

17	1257.5	39.7	554	2	Q35535	035535	rattus norv
18	1239.5	39.1	552	2	Q6P306	06P306	xenopus tro
19	1238	39.1	558	2	Q8K3R0	08K3R0	rattus norv
20	1232.5	38.9	558	2	Q8OZK3	08OZK3	mus musculu
21	1231.5	38.8	532	1	EST2_RABIT	P14943	oryctolagus
22	1230	38.8	559	1	EST2_HUMAN	U00748	homo sapien
23	1213.5	38.3	561	2	Q70631	070631	rattus norv
24	1213.5	38.3	566	2	Q766D7	0766D7	felis silve
25	1213.5	38.3	566	2	Q864S9	0864S9	felis silve
26	1210.5	38.2	557	2	Q640T6	0640T6	xenopus tro
27	1159	37.8	561	1	EST1_MESAU	Q64419	mesocricetu
28	1191	37.6	586	2	Q61PK9	061PK9	homo sapien
29	1187.5	37.5	568	2	Q6P674	06P674	mus musculu
30	1187.5	37.5	572	2	Q810S9	0810S9	mus musculu
31	1186.5	37.4	554	1	ESTM_MOUSE	Q95N05	canis fami
32	1183	37.3	565	2	Q95N05	Q95N05	canis fami
33	1179.5	37.2	571	2	Q6UW8	06UW8	homo sapien
34	1179	37.2	568	2	Q726J1	0726J1	homo sapien
35	1175	37.1	565	2	Q8TD29	08TD29	homo sapien
36	1171.5	37.0	568	2	Q8VCU1	08VCU1	mus musculu
37	1169.5	36.9	561	2	Q91WU0	091WU0	mus musculu
38	1164	36.7	565	2	Q97582	097582	sus scrofa
39	1162.5	36.7	555	2	Q6GMU1	06GMU1	brachydanto
40	1162.5	36.7	566	1	EST1_PIG	Q29550	sus scrofa
41	1159.5	36.6	566	2	Q46421	Q46421	macaca fasc
42	1150.5	36.3	565	1	EST1_RABIT	P12317	oryctolagus
43	1150	36.3	554	1	ESTN_MOUSE	P23953	mus musculu
44	1149	36.2	565	2	Q35534	Q35534	mesocricetu
45	1144	36.1	549	1	EST1_RAT	P10959	rattus norv

ALIGNMENTS

RESULT 1
ID Q6NT32 PRELIMINARY; PRT; 575 AA.
AC Q6NT32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE FLJ1547 protein.
GN Name=FLJ1547;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Butheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Sapichenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scherz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman A.J., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzanski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PCR rescued clones;

RA Director MGC Project;
RU Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL: BC069501, AAH69501.1; -.
DR HSSP: F12337; 1K4Y.
DR GO: GO:0003824; F: catalytic activity; IEA.
DR InterPro: IPR002018; Carboxylesterase.
DR InterPro: IPR000379; Ser esters.
DR Pfam: PF00135; Coesterase; 1.
DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE 575 AA; 63926 MW; 6F5B735BDFEC9C09 CRC64;

Alignment Scores:

Pred. No.:	2,25e-208	Length:	575
Score:	2915.00	Matches:	553
Percent Similarity:	97.54%	Conservative:	3
Best Local Similarity:	97.024	Mismatches:	6
Query Match:	91.964	Indels:	8
DB:	2	Gaps:	1

US-10-023-515-3 (1-1746) x Q6NTJ2 (1-575)

```
QY 34 TGGTCTTTTCTGATTTCTCGAGCCCTGTGGGACAGACAGAGTGGGAAAACTGGG 93
DB 14 TrrPAlaIletrrValIleuAlaIaPro-----ThrIySgLy 25
QY 94 CTTTCTGCGAAGGGGCGACAGAGAAACACAGGCTGGGATGAGTTGAGGGCGAAGCATG 153
DB 26 ProSerAlaGlnGlyProGlnArgAnthrArgLeuGlyTrrPLeGlnIySgLyVal 45
QY 154 ACTGTGCTGGGAAGCCCTGTGCTGTGAAGCTGTTCTCGAGTCCCTTGTGCTGCC 213
DB 46 ThrValIleuGlySerProValProValAsnValPheLeuGlyValProPheAlaIaPro 65
QY 214 CCGCTGGGATCCCTCGATTTACGAACCCGCAAGCTGATGCGCCCTGGGATTAATTGCGA 273
DB 66 ProLeuGlySerLeuArgPheThrAnPProGlnProIaSerProTrpAnPLeuArg 85
QY 274 GAAGCCACTCTTACCTTAATTTTGCTCCGAACTCGAGATGCTGTCTTAATCAA 333
DB 86 GlnAlaTrpSerTrpProAsnLeuCySLeuGlnAnSerGlyTrrPLeuLeuAspGln 105
QY 334 CACATGCTCAAGGTCATTAACCCGAAATTCGAGTGTGAGAGACTGCTTACCTGAC 393
DB 106 HisMetLeuIySValHisTrpProIySValSerGlnAspCySLeuTrpLeuAn 125
QY 394 ATCTATGCGCTGCGCCACGCGCATACAGGCTTCAAGCTCCCGCTTGGTGTGCCA 453
DB 126 IleTrpAlaProAlaHisAlaAspThrGlySerIySLeuProValIleuValTrpPhePro 145
QY 454 GGAGGTGCTTCAAGACTGAGCTCAGCTCCATCTTGTATGGGTCCGCCCTGGCTGCTAT 513
DB 146 GlyGlyAlaPheIySValHisTrpProIySValSerIlePheAspGlySerAlaLeuAlaIyTr 165
QY 514 GAGGACTGTGCTGTGCTGCTCCAGTACCGGCTAGAGAAATTTGGTTTCTTACACACA 573
DB 166 GlnAspValIleuValValValGlnTrpArgLeuGlyIlePheGlyPhePheThrThr 185
QY 574 TGGGATCGACATGCTCCCGGGAATGGGCTTCAAGACACAGGTGGCTGTCTGTCCG 633
DB 186 TrpAspGlnHisAlaIaProGlyAsnTrpAlaPheIySAspGlnValAlaAlaLeuSerTrp 205
QY 634 GTCCAGAAAGAACATGAGTCTTGGTGGGAGCCCACTCTGTGACATCTTTGGCGAG 693
DB 206 ValGlnIySAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGlyGln 225
QY 694 TCCGCGGAGGCATTAAGTGTCTTACTTACTGTCTCCATGGCGAAAGCTTATTC 753
DB 226 SerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaIyGlyLeuPhe 245
QY 754 CACAAAGCATCATGAGAGTGGGGTGGCCATCATCTTACCTTACCTGAGAGCCCATGATTTAT 813
```

```
DB 246 HisValAlaIleMetGlySerGlyValAlaIleIleProTrpLeuGlnAlaHisAspTrp 265
QY 814 GAGAAAGTGAAGACCTCGAGGTGGTGGCAATTTCTGTGTAGAAATGCTGAGACTCT 873
DB 266 GlnIySValSerGlnAspLeuGlnValValAlaHisPheCySValAsnAsnHisSerAspSer 285
QY 874 GAGGCGCTCTGAGGTGCTTGAAGACAAACCTTCAGAGAGCTGTGACCTCAGCCAG 933
DB 286 GlnAlaIleuAlaArgCySLeuArgThrIySProSerIySLeuLeuThrIleuSerGln 305
QY 934 AAAACAAAGCTTTTCACTCGAGGTGAGTGGTGGCTTTCTTCTTAAGGCTCTAGAT 993
DB 306 IySThrIySValSerPheThrArgValAlaAspGlyAlaPhePheProAsnGlnProLeuAsp 325
QY 994 CTATTGTCTGAGAAAGCATTTTAAGCAATTCCTTCATCATGCGAGTCAATTAAGCAGAG 1053
DB 326 LeuLeuSerGlnIySAlaIlePheIySAlaIleProSerIleIleGlyValAsnAsnHisGln 345
QY 1054 TGTGGCTTCCCTGCTGCTTATGAGAGAGGCTCTGAGATCTTGAAGCTCCCAAGATCC 1113
DB 346 CySgIyPheLeuLeuProMetIySgIyAlaProGlnIleLeuSerGlySerAnIySValSer 365
QY 1114 CTTGCGCTCCATCTGATACAAACATCTCGACATCCGCGCTCAGTATTTGGACCTTGTG 1173
DB 366 LeuAlaIleuHisIleuIleGlnAsnIleLeuHisIleProProGlnIySLeuHisIleuVal 385
QY 1174 GGTAAATGAATCTTCATGACACAGCATCTCTGATGAAATCGAGACAGTCTTGGAC 1233
DB 386 AlaAsnGlyIyTrpHisIleAspIySValSerLeuThrGlnIleArgAspSerIleuAsp 405
QY 1234 TTTGCTTGAAGATGAGTCTTTTGTGGCTCCGACATGACAGAGCTCGATACAGAGAT 1293
DB 406 LeuLeuGlyAspValPhePheValAlaProAlaLeuIleTrpAlaArgTrpHisIleArgAsp 425
QY 1294 GCTGTGACACTGTCTACTTCTATGAGTTTGGGACCGGCTCAGTGTGTTGAAGACAG 1353
DB 426 AlaGlyAlaProValIyTrpPheTrpGlnPheArgHisArgProGlnCySValPheGlnAspThr 445
QY 1354 AAGCGGCTTTTGTGCAAGCGGACGACGCTGATGAAGTCCGCTTGTGTGGTGTGCC 1413
DB 446 IySProAlaPheValIySValAspHisIleAlaAspGlnValAlaArgPheValPheGlyGlyAla 465
QY 1414 TTCTGAAAGGGGAGCATGTTATGTTGGAAGAGACGAGGAGGAGAAATTAATCTGAGC 1473
DB 466 PheLeuIySgIyAspIleValIleMetPheGlyIyAlaThrGlnGlnGlyIySLeuLeuSer 485
QY 1474 CGGAAGATGATAAATACCTGGGCTAACCTTTGCTGGAACCGGGAATCTTAATGGGAACGAC 1533
DB 486 ArgIySValMetClyIyTrpAlaIleThrPheAlaArgThrGlyAnPProAsnGlyAsnAsp 505
QY 1534 CTGTCTCTGTGGCGAGCTTAAATCTGATGACAGAGTACCTTCAGCTGGAATTTGAACATG 1593
DB 506 LeuSerLeuTrpProAlaIyTrpAsnLeuThrGlnIyTrpLeuGlnIleAspLeuAsnMet 525
QY 1594 AGCTCGGACAGAGACTCAAAAGACCGGAGGTGAGTTTGTGACAGAGACATGCCCTGTG 1653
DB 526 SerIleuGlyGlnArgLeuIySgIyProArgValAspPheTrpHisSerThrIleProIleu 545
QY 1654 ATCTGTCTGCTCGACATGCTCCACAGTCTCTTTCTTCTTAATCTTCTCTCTCTC 1713
DB 546 IleLeuSerAlaSerAspMetLeuHisSerProLeuSerSerIleuThrPheLeuSerIleu 565
QY 1714 CTCAGGCTTCTTTCTTTCTTTGGCTCCT 1743
DB 566 LeuGlnProPhePhePheCysAlaPro 575
```

RESULT 2

ID	Q96DN9	PRELIMINARY:	PRF:	525	AA.
AC	Q96DN9:				
DT	01-DEC-2001	(TrEMBLrel. 19, Created)			
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)			


```
Db 266 G|U|Y|S|E|G|I|A|B|P|E|U|I|N|V|A|I|A|I|A|H|E|P|H|E|C|Y|G|I|Y|A|N|B|A|N|A|S|E|A|B|S|E| 285
QY 874 G|A|G|G|C|C|T|G|T|A|G|G|G|C|C|T|G|A|G|A|C|A|A|A|C|C|T|C|A|A|G|A|C|T|G|T|G|A|C|C|G|C|G| 933
Db 286 G|U|A|A|L|E|U|A|A|G|Y|E|U|A|I|G|T|H|Y|S|P|R|O|S|E|L|Y|S|G|I|U|E|U|E|U|H|E|U|H|E|U|S|E|G|I|N| 305
QY 934 A|A|A|A|A|A|G|T|C|T|T|T|C|A|C|T|G|A|G|G|T|G|A|T|G|G|T|G|T|T|T|C|T|T|A|T|G|A|G|C|C|T|C|A|T| 993
Db 306 L|Y|S|T|H|Y|S|E|R|P|H|E|T|H|R|A|G|V|A|I|A|A|P|R|G|I|A|A|P|H|E|R|P|R|O|A|N|G|I|U|R|P|R|O|L|E|U|A|P|R 325
QY 994 C|T|A|T|G|T|C|T|C|A|G|A|A|G|C|A|T|T|T|A|A|G|C|A|T|T|C|T|C|A|T|C|A|T|G|A|G|A|G|C|A|T|A|C|C|A|G| 1053
Db 326 L|E|U|E|U|S|E|G|I|N|L|Y|A|I|A|P|H|E|Y|S|A|I|A|I|E|P|R|O|S|E|R|I|E|I|G|I|Y|A|A|N|B|A|N|H|I|S|G|I| 345
QY 1054 T|G|T|G|G|C|T|C|G|T|C|C|C|T|A|T|G|A|A|G|A|G|G|C|C|T|G|A|G|A|G|C|C|T|G|A|G|G|C|C|C|A|A|A|G|T|C| 1113
Db 346 C|Y|S|G|I|Y|P|H|E|L|E|U|H|E|R|P|R|O|K|E|T|Y|S|G|I|U|A|P|R|O|G|I|U|I|E|U|S|E|R|G|I|Y|S|E|R|A|N|L|Y|S|E|R 365
QY 1114 C|T|T|G|C|C|T|C|C|A|T|C|T|G|A|T|A|C|A|A|A|A|C|A|T|C|T|G|C|A|C|A|T|C|C|G|C|C|C|A|G|A|T|T|T|G|C|A|C|T|T|G| 1173
Db 366 L|E|U|A|A|L|E|U|H|I|S|E|U|I|I|E|G|I|N|B|A|N|I|E|U|H|I|S|I|E|R|P|R|O|G|I|N|T|Y|L|E|U|H|I|S|E|U|A|I| 385
QY 1174 G|C|T|A|T|A|A|A|T|A|C|T|C|A|T|G|A|C|A|A|G|C|A|C|T|C|C|T|G|A|T|G|A|A|T|C|G|A|G|A|C|A|G|T|T|T|G|A|C| 1233
Db 386 A|I|A|A|N|G|I|U|Y|R|H|E|H|I|A|B|R|Y|S|H|I|S|E|R|L|E|U|H|I|G|I|U|I|E|A|R|G|A|B|S|E|R|L|E|U|A|B|R 405
QY 1234 T|T|G|C|T|G|G|A|G|A|T|G|T|T|T|G|T|G|G|C|C|T|G|C|A|G|A|C|A|G|C|C|G|C|G|A|T|A|T|C|A|G|A|G|A|T| 1293
Db 406 L|E|U|E|U|S|G|I|A|B|R|V|A|A|P|H|E|R|H|E|V|A|I|A|P|R|O|A|L|E|U|I|I|E|T|H|A|I|A|R|G|Y|H|I|A|R|G|--- 424
QY 1294 G|C|T|G|T|G|A|C|C|T|G|T|C|T|A|C|T|T|T|A|G|T|T|C|G|G|C|A|C|C|G|C|C|T|C|A|G|T|G|C|T|T|T|G|A|A|G|A|C|A|G| 1353
Db 424 ----- 424
QY 1354 A|A|G|C|G|C|G|T|T|T|G|T|G|A|A|G|C|G|C|A|C|G|C|T|G|A|A|G|T|C|C|G|T|T|G|T|G|G|T|G|G|T|G|C| 1413
Db 424 ----- 424
QY 1414 T|T|C|T|G|A|A|G|G|G|G|G|A|C|A|T|T|T|A|T|G|T|T|G|A|A|G|G|G|C|A|G|G|A|G|G|A|G|A|G|A|G|T|T|A|C|T|G|A|G|C| 1473
Db 425 ----- -G|U|G|Y|A|A|T|H|I|G|I|U|G|I|U|Y|L|E|U|E|U|S|E|R 435
QY 1474 C|G|A|A|G|A|T|G|A|T|G|A|A|T|A|C|T|G|G|G|C|A|C|T|T|T|G|T|G|C|G|A|C|C|G|G|A|A|T|C|T|A|T|T|G|G|A|C|G|A|C| 1533
Db 436 A|R|G|Y|S|E|T|E|U|C|E|T|Y|T|R|P|A|A|H|R|H|E|A|I|A|A|G|T|H|I|G|Y|A|S|R|P|R|O|A|N|G|I|Y|A|N|B|R 455
QY 1534 C|T|G|T|C|T|G|T|G|G|C|G|C|T|T|A|T|A|T|G|A|C|T|G|A|C|G|T|A|C|T|C|C|A|G|C|T|T|G|A|C|T|T|G|A|C|A|T|G| 1593
Db 456 L|E|U|S|E|R|L|E|U|T|R|P|R|O|A|I|T|Y|R|A|N|L|E|U|H|R|G|I|U|G|I|N|T|Y|L|E|U|G|I|N|L|E|U|A|B|R|E|U|A|N|E|T 475
QY 1594 A|G|C|C|T|G|G|A|G|A|G|A|C|T|A|A|A|A|A|C|G|G|G|G|G|G|A|T|T|T|T|G|G|A|C|A|G|A|C|A|T|C|C|C|C|T|G| 1653
Db 476 S|E|R|E|U|G|I|Y|A|N|B|Y|E|U|S|I|U|R|A|G|V|A|A|P|R|H|E|R|P|R|H|S|E|R|T|H|R|I|E|R|P|R|O|L|E|U 495
QY 1654 A|T|C|T|G|T|G|C|C|T|G|C|A|G|A|T|G|C|C|A|G|A|G|C|C|T|T|T|C|T|T|A|C|T|T|A|C|T|T|C|T|C|T|C| 1713
Db 496 I|L|E|U|S|E|R|A|I|A|S|E|R|A|S|P|R|E|U|H|I|S|E|R|P|R|O|L|E|U|S|E|R|S|E|R|L|E|U|H|T|H|R|P|H|E|U|S|E|R|I|E|U 515
QY 1714 C|T|C|A|G|C|T|T|T|T|T|T|T|T|T|T|T|G|T|G|T|C|C|T 1743
Db 516 L|E|U|G|I|N|P|R|O|A|H|P|H|E|R|P|H|E|C|Y|S|A|A|P|R|O 525
```

```
RESULT 3
Q8NBC8 PRELIMINARY; PRT; 469 AA.
AC Q8NBC8;
DT 01-OCT-2002 (TRENBLREL. 22, Created)
DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE Hypothetical protein FLJ33678.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakematsu A., Hayashi K., Sato H., Nagai K., Kiyama K., Makita H.,
RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishi S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahata K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoliti T., Kaku Y., Kodaira H., Kondoh H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa M., Ozaki Y.,
RA Abe K., Kamihara K., Katsura N., Saito K., Tanikawa M., Yamazaki K.,
RA Niwomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tani H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotaka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Komiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki M., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okutani R., Kawakami T., Noguchi S., Itoh T., Shigetani K., Senba T.,
RA Matsunuma K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RL Nat. Genet. 36:40-45(2004).
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AK090997; BAC03565.1; -.
DR HSSP; P12337; IKXY.
DR GO; GO:0016787; F:Hydrolase activity; IEA.
DR InterPro; IPR002018; CarboxylesteraseB.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KM Hydrolase.
SQ SEQUENCE 469 AA; 52312 MW; 7A188F68C10A9080 CRC64;
Alignment Scores:
Pred. No.: 4,72e-174 Length: 469
Score: 2454.00 Matches: 2
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.57% Mismatches: 0
Query Match: 77.41% Indels: 0
DB: Gaps: 0
US-10-023-515-3 (1-1746) x Q8NBC8 (1-469)
QY 337 A|T|G|C|T|C|A|G|G|T|G|C|A|T|T|A|C|C|G|A|A|T|T|G|G|A|G|T|G|C|A|G|A|G|A|C|T|G|C|T|A|C|T|G|A|C|A|T|C| 396
Db 1 M|E|U|E|U|S|V|A|I|H|I|S|T|Y|R|P|R|O|L|Y|S|P|H|E|G|I|Y|A|I|S|E|R|I|U|A|B|R|C|Y|S|E|R|Y|L|E|U|A|N|I|E 20
QY 397 T|A|T|G|C|C|T|G|C|C|C|A|C|G|C|G|C|A|T|A|C|G|G|T|C|C|A|G|G|T|C|C|C|C|C|T|T|G|T|G|T|G|T|T|C|C|C|A|G|A 456
Db 21 T|Y|R|A|P|R|O|A|H|A|B|A|A|P|R|H|G|Y|S|E|R|Y|S|E|R|V|A|L|E|U|A|I|T|R|P|H|E|R|P|R|O|G|I| 40
QY 457 G|G|T|G|C|T|T|C|A|G|A|C|T|G|G|C|T|C|A|G|C|C|T|C|A|T|G|A|T|G|G|T|G|G|C|C|G|C|G|G|C|T|G|C|T|A|T|G|A|G| 516
Db 41 G|I|Y|A|A|H|E|Y|S|T|R|G|I|Y|S|E|R|A|I|A|S|E|R|I|E|P|H|E|A|S|P|I|Y|S|E|R|A|L|E|U|A|I|A|I|T|Y|G|I|U 60
QY 517 G|A|G|C|T|G|T|G|T|G|T|G|T|G|T|G|C|C|A|G|T|A|C|G|G|C|T|A|G|A|A|T|A|T|T|G|G|T|T|C|T|C|A|C|A|C|A|T|G| 576
```

```

Db      61 AspValIeuValIValIValGlnTyArgLeuGlyIlePheGlyPhePheThrThrTrp 80
Qy      577 GATAGACATGCTCGGGGACATGGGCTTCAAGACACAGGGGCTGTGCTGGGCT 636
Db      81 AspGlnHisAlaProGlyAsnTrpAlaPheLeuAspGlnValAlaAlaLeuSerTrpVal 100
Qy      637 CAGAGAACATCGAGTCTTGGTGGGAGCCCGACCTGTGACCATCTTTGGCAGATCC 696
Db      101 GlnTyAsnIleGlnPhePheGlyGlyAspProSerSerValThrIlePheGlyGlySer 120
Qy      697 CGGGAGCCATTAAGTCTTACTTACTTACTTCTCCCATGAGCCAAAGGCTTATCCAC 756
Db      121 AlaGlyAlaIleSerValSerSerLeuIleLeuSerProMetAlaGlyGlyLeuPheHis 140
Qy      757 AAAGCATCATGAGAGTGGGGTGGCATCATCCCTTCCCTGGAGGCCCATATATATAG 816
Db      141 LysAlaIleMetGlnSerGlyValAlaIleIleProTyLeuGlnAlaHisAspTrpGlu 160
Qy      817 AAGAGTGAAGACCTGACAGTGGTGGTGCATTTCTGTGGTAACAATGCGTCAAGACTGAG 876
Db      161 LysSerGlnAspLeuGlnValValAlaHisPheGlyGlyAsnAsnAlaSerAspSerGlu 180
Qy      877 GCCCTGCTAGAGTGCCTGAGACAAACCCCTCAAGAGCTGCTGACCTCAGCCAGAA 936
Db      181 AlaLeuLeuArgCysLeuAlaGThrIysProSerIysGlnLeuLeuThrLeuSerGlnIys 200
Qy      937 ACAAGTCTTTCATCTCGAGTGGTGTATGTGCTTCTTCCAAAGACCTCTATATCTA 996
Db      201 ThrIysSerPheThrArgValValAspGlyAlaPhePheProAspGlnProLeuAspLeu 220
Qy      997 TTGTCTCGAAGACATTTAAAGCAATTCCTTCATCATCGAGTCAATCAACACAGAGTGT 1056
Db      221 LeuSerGlnLysAlaPheLeuAlaIleProSerIleIleGlyValAsnSerHisGlyCys 240
Qy      1057 GCGTCTCTGCTGCTATGAGAGGCTCTGAGATCTTCACTGAGCTCAACAAGTCCCTT 1116
Db      241 GlyPheLeuLeuProMetIleGlnLysAlaProGlnIleLeuSerGlySerAsnLysSerLeu 260
Qy      1117 GCCCTCCATCTGATCAAAACATCTCTGCACATCCCGCTCAGTATTTGCACTTGTGCT 1176
Db      261 AlaLeuHisLeuIleGlnAsnIleLeuHisIleProProGlnIleLeuHisLeuValAla 280
Qy      1177 AATGAATACTCTCATGACACAGACCTCCCTGACTGAATCCGAGACAGTCTTGGACTTG 1236
Db      281 AsnGlnTyArgPheHisAspHisSerLeuThrGlnIleHisAspSerLeuLeuAspLeu 300
Qy      1237 CTTGAGATGTGTTCTTGTGTGCTCTGCTGACATCACAGCTGCATATCACAGAGTGT 1296
Db      301 LeuGlyAspValAlaPhePheValValProAlaLeuIleThrAlaArgTyHisArgAspAla 320
Qy      1297 GGTGACCTGTCTACTTCTATGAGTTTGGGACCGGGCTCAGTCTTTGGAAGACAGAA 1356
Db      321 GlyAlaProValTyArgPheTyGlnPheArgHisAspProGlnCysPheGlnAspThrIys 340
Qy      1357 CGGGCTTTTTCAGAACGCGACGAGCTGAAGTCCGCTTGTGTTGGTGGTGGCTTTC 1416
Db      341 ProAlaPheValLysAlaAspHisAlaAspGlnValHisArgPheValPheGlyGlyAlaPhe 360
Qy      1417 CTGAAGGGGACATTTGTTATGTTCTCGAAGAGCCAGAGAGAGAGAGTACTGAGCCGG 1476
Db      361 LeuTySerGlyAspIleValMetPheGlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArg 380
Qy      1477 AAGATGATGAATACTGGGCTTACCTTGTGCTGACCCGGGAATCTTAATGGGAACAGCTG 1536
Db      381 LysMetMetLysTyArgAlaThrPheAlaArgTrpGlyAsnProAsnGlyAsnAspLeu 400
Qy      1537 TCTCTGAGGCGAGCTTATATCTGACTGACAGTCACTCCAGCTGGAAGCTTGAACATGAGC 1596
Db      401 SerLeuTrpProAlaTyArgAsnLeuThrGlnGlnTyLeuGlnLeuAspLeuAsnMetSer 420
Qy      1597 CTCGACACAGACCTCAAGAACCCGGGGTGAATTTTGGACAGACCATCCCTGATC 1656
Db      421 LeuGlnGlnArgLeuLeuGlnLysProArgValGlnPheTrpThrSerThrIleProLeuIle 440

```

```

Qy      1657 CTGTCTGCTCCGACATGCTCCACAGTCTCTTCTTCTTAACTTCTCTCTCTC 1716
Db      441 LeuSerAlaSerAspMetLeuHisSerProLeuSerSerLeuThrPheLeuSerLeuLeu 460
Qy      1717 CAGCTTCTCTTCTTCTTCTTGTGCTCT 1743
Db      461 GlnProPhePhePheCysAlaPro 469

RESULT 4
ID      06AM47      PRELIMINARY;      PRT;      575 AA.
AC      06AM47;
DT      25-OCT-2004 (TREMBLrel. 28, Created)
DT      25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT      25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE      Carboxylesterase-like urinary excreted protein.
GN      Name=cauxin;
OS      Canis familiaris (Dog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX      NCBI_TaxId=9615;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Miyazaki M., Yamashita T., Taira H., Suzuki A.;
RT      "cauxin family protein."
RL      Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
CC      -1. SIMILARITY: Belongs to the type-B carboxylesterase/11pase family.
EMBL; AB186392; BAD35015.1; -.
GO; GO:0003824; F: catalytic activity; IEA.
DR      InterPro; IPR002018; CarboxylesteraseB.
DR      InterPro; IPR000379; Ser_ester.
DR      Pfam; PF00135; Coesterase; 1.
DR      PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR      PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW      Hydrolase.
SQ      SEQUENCE 575 AA; 63620 MW; 1005C35B82E1183D CRC64;

Alignment Scores:
Pred. No.:      8.09e-159      Length:      575
Score:      2250.00      Matches:      417
Percent Similarity:      86.75%      Conservative:      61
Best Local Similarity:      75.68%      Mismatches:      73
Query Match:      70.98%      Indels:      0
DB:      2      Gaps:      0

US-10-023-515-3 (1-1746) x 06AM47 (1-575)
Qy      91 GGGCTTCTGCTGAAGGGCCACAGAGAACACAGGCTGGATGATTCAAGGCAAGCA 150
Db      25 GlyProAlaHisAlaAspAlaProValArgSerThrArgLeuGlyTrpValArgGlyLysGln 44
Qy      151 GTCACTGTGCTGGAGAGCCGTGTGCTGTGAACGTGTTCTTGGAGTCCCTTGTGCT 210
Db      45 AlaThrValLeuGlySerThrMetProValAsnValPheLeuGlyIleProPheAlaAla 64
Qy      211 CCCCCGTGGGATCCCTCGATTTACGAACCGGACGCTGCATGCCCTGGGATAACTTG 270
Db      65 ProProLeuGlyProLeuAlaArgPheLysArgProLysProAlaLeuLeuThrAsnAspSer 84
Qy      271 CGAGAGACCACTCTTACCTTAATTTGTGCTCCGAGACTCAGAGTGGCTGCTTGAAT 330
Db      85 ArgAspAlaHisSerTyArgLysLeuGlySerGlnHisAsnSerValTrpLeuLeuSerAsp 104
Qy      331 CAACATGCTCAAGTCAATTAACCGAAATTCGAGAGTGCAGAAAGTGCCTTAACCTG 390
Db      105 GlnHisPheLeuLysValHisTyArgProAsnLeuGlnValSerGlnAspCysLeuTyTrpLeu 124
Qy      391 AACATCTATGGGCTGGCCCAAGCCGATACAGGCTTCAAGCTCCCGCTGTGTGTGCTTC 450
Db      125 AsnIleTyArgAlaProAlaHisAlaAsnThrGlySerLysLeuProValMetValTrpPhe 144
Qy      451 CCAGAGAGTGCCTTCAAGACGTGCTCAGCTTCATCTTGAATGGTGCAGCCCTGAGCTGCC 510

```



```

334 CACATGCTCAAGTGCATTTACCCGAAATTCGGAGTGTCAAGAACTGCTCTTACCTGAC 393
      |||||
106 AasnleuLeuValSerProIleleuGluMetSerGluAspCyleuLeuYtrleuAen 125
      |||||
394 ATCTATGGGCTGGCCCAAGCCGATACAGGCTCAAGCTCCCGCTTGGTGTGTTCCCA 453
      |||||
126 IletYtrAlaProCysHlaAlaAbaNbnGlySerSerLeuProValMetAlaTrpIlePro 145
      |||||
454 GGAGGCTGCTTCAAAAGTGGCTCAGCTCTCCATCTTGAATGGGTCCGCGCTGCTGCTAT 513
      |||||
146 GtlyGtlyGtlyPheGtlyThrGtlySerAlaSerIlePheAspGlySerAlaLeuAlaValYtr 165
      |||||
514 GAGAGCTGCTGTGTGCTGCTCCAGTACCGGCTAGAAATATTTGTTTCTTCAACA 573
      |||||
166 GluAspValIleValValThrIleGtlyTrpArgLeuGtlyIlePheGtlyPhePheThrThr 185
      |||||
574 TGGGATCAGCATGCTCCGGGAACTGGGCTTCAAGAACAGAGTGGCTGTGCTTGG 633
      |||||
186 GluAsnGlnHlaAlaProGtlyAbaNtrPalaPheTrpAspGlnLeuAlaIleLeuLeuTrp 205
      |||||
634 GTCCAGAAAGATCAGTGTCTTGGTGGGAGCCGAGCTGTGACCACTTTGGCGAG 693
      |||||
206 ValArgGluAbaNtrIleYtrPheGtlyGtlyAbaNtrPalaPheSerValThrIlePheGtlyAba 225
      |||||
694 TCCGCGGAGGCAATAGTGTCTTATCTTATCTGTCTCCAGTGGCCAAAGCTTATTC 753
      |||||
226 SerAlaGtlyAlaIleSerIleSerSerLeuIleLeuSerProLeuSerAlaAspLeuPhe 245
      |||||
754 CACAAAGCCATCAGAGAGTGGGCTGGCCATCATCTCTTACCTGAGGCCCATGATAT 813
      |||||
246 HisArgAlaIleMetGlnSerGtlyAlaAlaIleIleProSerLeuYtrSerSerAspAba 265
      |||||
814 GAGAAAGATGAGAGCTGAGTGTGTCACATTTCTGTGGTAAACATCCGTCACATCT 873
      |||||
266 AspLeuYtrHlaAspLeuGlnAlaValAlaAbaNValCysAspCysAbaNValSerAspSer 285
      |||||
874 GAGGCTGCTGAGAGTGGCTGAGACAAACCCCTCAAGAGACTGTGACCTTCAAGCCAG 933
      |||||
286 LysAlaIleLeuLeuYtrCyleuAtrGtlyYtrSerSerLeuGluIleMetSerLeuSerGln 305
      |||||
934 AAAACAAAGCTTCTTACCTGAGTGTGATGAGTGTCTTTCTTCTTAAAGCCCTTGAAT 993
      |||||
306 LysAlaIleYtrSerPheThrArgValAlaAspGlySerPhePheSerGtlyGluProLeuGln 325
      |||||
994 CTATGTGCTCAGAAAGATTTAAAGCATCTCTTCAATCATGAGAGTCAATTAACACAG 1053
      |||||
326 LeuLeuSerGlnYtrThrLeuYtrIleValProSerIleIleGtlyValAbaNbnGlnGln 345
      |||||
1054 TGTGCTTCCGCTGCTGATGAAAGAGGCTCTGAGATCTCAGTGGCTCAACAGTCC 1113
      |||||
346 CysGtlyYtrIleLeuProValAtrAspThrProGtlyIleLeuGtlySerAbaNbnGlnSer 365
      |||||
1114 CTGGCCCTCAGTGTGATCAAAACATCTGCAATCCGCGCTCAGATTTTGGACCTTGTG 1173
      |||||
366 ThrAlaLeuThrIleIleHlaThrLeuLeuHlaIleProThrGlnHlaIleLeuYtrIleVal 385
      |||||
1174 GCTAATGAATCTCTTCAATGACAAAGCATCTGCTGATGAATCCGAGACAGTCTTGTGAC 1233
      |||||
386 ThrIleYtrGtlyYtrPheHlaGtlyYtrHlaSerProIleAspIleArgAspThrThrLeuAba 405
      |||||
1234 TTGCTTGGAGATGATGTTCTTGTGCTGCTGCTGATGACAGTCAAGTCAAGAT 1293
      |||||
406 LeuPheGtlyAspValAlaPhePheValAlaProGtlyLeuValAlaIleArgTyrHlaAspArg 425
      |||||
1294 GCTGTGACAGCTGTCTTCTTATGATGATTTGGACCCGCGCTCAGTGTCTTGAAGACAG 1353
      |||||
426 SerGtlyGtlyProValYtrPheYtrGtlyPheGlnHlaAspProHlaCysPheGtlyAbaNbnSer 445
      |||||
1354 AAGCGGCTTTTGTCAAGCCGACCAAGCTGATGAAAGTCCGCTTGTGTGCTGCTGCTGCC 1413
      |||||
446 ArgProAlaAlaPheValIleValAlaAspHlaThrArgGtlyIleArgPheValAlaPheGtlyPro 465
      |||||
1414 TTCTGAAAGGGGAGCATTTGTTATGTTCAAGAGACCAAGAGAGAGAAAGTTACTGAGC 1473

```

```

Db 466 PheLeuYtrGtlyAspValAlaMetPheGtlyGtlyAlaThrGtlyGtlyLeuLeuSer 485
      |||||
Qy 1474 CGGAAGATGATGAATTACTGGGCTACCTTGTGTCGAACCGGAGATCTTAATGGGAACGAC 1533
      |||||
Db 466 ArgGlyMetMetLeuYtrTrpAlaAbaNbnPheAlaAspSerGtlyAspProAbaNbnIleAsp 505
      |||||
Qy 1534 CTGTCTGTGCTGCGCAGCTTATATCTGACTGACGATCACTTCCAGCTGATCTTGAACATG 1593
      |||||
Db 506 LeuProProTrpProValYtrAspGluAbaNbnGtlyGtlyLeuGtlyAbaNbnIle 525
      |||||
Qy 1594 AGCTTCGACAGACAGTCAAGAACCCGCGGTGATTTTGGACGACCACTCCCGCTG 1653
      |||||
Db 526 SerThrGtlyArgArgLeuYtrAspGlnArgValGtlyPheTrpThrAspThrLeuProLeu 545
      |||||
Qy 1654 ATCTGTGTGCTCCGACATGTCACAGTCTCTTCTTCTTAACTTCTCTCTCTC 1713
      |||||
Db 546 IleLeuSerAlaSerIleAlaLeuLeuSerProThrPheSerLeuIleLeuLeuSerLeu 565
      |||||
Qy 1714 CTCCAGCTTCTTCTTCTTCTTGTGCT 1740
      |||||
Db 566 LeuSerProValLeuLeuSerAlaAla 574
      |||||

RESULT 6
ID 081034 PRELIMINARY; PRT: 545 AA.
AC 081034:
DT 01-MAR-2003 (TRIMBLREL. 23, Created)
DT 01-MAR-2003 (TRIMBLREL. 23, Last sequence update)
DT 01-OCT-2003 (TRIMBLREL. 25, Last annotation update)
DE Carboxylesterase-like urinary excreted protein.
GN Name=Caxlin;
OS Fells silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Fells.
OX NCB1_TaxID=9685;
RN [1]
RP TISSUE=KIDNEY;
RC MEDLINE=22458314; PubMed=12401131; DOI=10.1042/BJ20021446;
RX Miyazaki M., Kamile K., Soeta S., Taira H., Yamashita T.;
RT "Molecular cloning and characterization of a novel carboxylesterase-
RT like protein that is physiologically present at high concentrations in
RT the urine of domestic cats (Fells catus).";
RL Biochem. J. 370:101-110(2003).
CC -!- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AB045377; BAC22577.1; -.
DR HSSP; P12337; 1K4Y.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002018; CarbesteraseB.
DR InterPro; IPR000379; Ser_estrb.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
SQ SEQUENCE 545 AA; 60505 MW; 9F73FA69D271FA9 CRC64;

Alignment Scores:
Pred. No.: 4,44e-139 Length: 545
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 62.62% Indels: 2
DB: Gaps: 2

US-10-023-515-3 (1-1746) x 081034 (1-545)
Qy 91 GGGCTTGTGCTGAAAGGGCCACAGAGAAACAGAGCTGGATGATTCAGGGAAGCA 150
      |||||
Db 25 GlyProAlaAlaAspAlaProValAtrGserThrArgLeuGtlyTrpValArgGtlyYtrGln 44
      |||||
Qy 151 GTCACTGTGCTGGGAAGCCGTGCTGTGAACGTGTTCTTCAAGTCCCTTTCGTCT 210
      |||||

```

```

Db      45  ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyIleProTyrlaIala 64
Qy      211  CCCCCTGGGAGATCCCTGGATTTAGAAACCCGAGACCTGATCCGCCCTGGATAACTGG 270
Db      65  ProPheLeuGlyProLeuArgPheLeuGlnProLysProIalaLeuProLysAsnPhe 84
Qy      271  CGAAGACCACTCTCTACCTTAATTTGCTCCAGAACTCAAGAGTGGCTGCTTAGAT 330
Db      85  ArgAsnAlaThrSerTyProLysLeuCyPheGlnAspLeuGlnTrpLeuValSerTyx 104
Qy      331  CAACACATGCTCAAGTGTGATTAACCGAAATTCGAGAGTCAAGAACTGCTTACCTG 390
Db      105  GlnHsValLeuLeuValArgTyProLysLeuGlnAlaSerGlnAspCyLeuTyx 124
Qy      391  AACATCTATGCGCTGCGCCAGCCGATACAGGCTCCAGGCTCCGCTGCTGGTGGTTC 450
Db      125  AsnIleTyrlaProIalaHsAlaAspAsnGlySerAsnLeuProIalaMetValTrpPhe 144
Qy      451  CCAGAGAGTGCCTTCAAGACTGGCTCAAGCTTCATTTGATGGGTCCGCGCTGGCTGCC 510
Db      145  ProGlyGlyAlaPheLeuMetGlySerAlaSerSerPheAspGlySerAlaLeuAlaIa 164
Qy      511  TATGAGACGCTGCTGCTGCTGCTCCAGTACCGGCTTGAATATTTGGTTTCTTCAAC 570
Db      165  TyrgIuaAspValLeuIleValThrThrGlnTyArgLeuGlyIlePheGlyPhePheAsp 184
Qy      571  ACATGGATCAGCATGCTCCGGGGAATGGGCTTCAAGGACAGAGTGGCTGCTGTC 630
Db      185  ThrGlyAspGlnHsAlaArgIlyAsnTrpAlaLeuLeuAspGlnValAlaIalaLeuThr 204
Qy      631  TGGGTCAGAAAGAACATGAGTCTTCCGTTGGTGGGACCCGAGCTGTGAGCAATCTTTGGC 690
Db      205  TrpValArgAspAsnIleGluPhePheGlyGlyAspProArgSerValThrIlePheGly 224
Qy      691  GAGTCCGGGGAGCCATAGTGTCTTAGCTTAACTGCTTACCTGCCATGGCCAAAGCTTA 750
Db      225  GluSerAlaGlyAlaIleSerAlaSerSerLeuIleLeuSerProIalaAsnIlyLeu 244
Qy      751  TTCCCAAAAGCCATCATGAGAGTGGGAGTGGGCTCATCCCTTACCTG---GAGGCCCAT 807
Db      245  PheHsIlyValAlaIleMetGluSerGlyValAlaIleLeuProLeuLeuMetArgProPro 264
Qy      808  GATTATGAAAGAGTGAAGACCTTGAGGCTGTGCAATTTCTGTGTGAACATGCTGCA 867
Db      265  GlyAspGlyArgLysLysAspLeuGlnValIleAlaArgIleCyGlyGlyHsAlaSer 284
Qy      868  GACCTGAGGCGCTGCTGAGGCTGAGCAAAACCTCCAGAGACTGCTGACCTC 927
Db      285  AspSerAlaAlaLeuLeuGlnCyLeuArgAlaLysProSerGlnGlnLeuMetAspIle 304
Qy      928  AGCCAGAAACAAAGTCTTTGACTCGAGTGGTGAATGGTCTTTCTTCTTAATGAGCCT 987
Db      305  SerIlyLysLeuThrPheSerIleProValIleAspAspPhePheProAspGluPro 324
Qy      988  CTAAGATCTATGCTCAGAAAGCATTTAAAGCAATTTCTTCATCATGCGAGTCAATAC 1047
Db      325  ValAlaLeuLeuTrpGlnLysAlaPheAsnSerValProSerIleIleGlyValAlaAsn 344
Qy      1048  CACAGAGTGGCTTCCGCTGCTGCTGTAAGAGAGGCTCTGAGTCTCCAGTGGCTCAAC 1107
Db      345  HisGlnCyValaPheLeuLeu---SerThrGluPheSerGlnIleLeuGlyIleSerAsn 363
Qy      1108  AAGTCCCTGCTGCTCATGATGATCAAAACATCTGTCATCCGCTCAGATATTTGAC 1167
Db      364  ArgSerLeuAlaLeuTyrlaLeuValHsIleThrPheLeuAsnIleProTrpGlnTyrlaHs 383
Qy      1168  CTTTGCTTAATGAATACCTTCAATGACAGCACTCCCTGACTGAATTCGAGACAGTCTT 1227
Db      384  LeuValAlaAspHsAlaTyrlaTyrlaAsnLysHsSerProValGlnIleArgAspSerPhe 403
Qy      1228  CTGAGCTTGGTGGAGATGGTGTCTTTGGTGGCTGCTGACATCCGCTCAGATATTCAC 1287
Db      404  LeuAspLeuLeuGlyAspValLeuPheValValProGlyValValThrAlaArgTyrlaHs 423

```

```

Qy      1288  AGAGAATGCTGTCGACCTGTCTACTTTATGAGTTTCGCGACCGGCTCACTGCTTTGAA 1347
Db      424  ArgAspAlaGlyAlaProValTyrlaPheTyrlaGluPheGlnHsProProGlnCyLeuAsn 443
Qy      1348  GACAGAAAGCGGGCTTTTGTCAAAAGCCGACCAAGCTGATGAAAGTCCGCTTGTGTCGT 1407
Db      444  AspThrArgProAlaPheValAlaPheAlaAspHsIleSerAspGlnIleArgPheValPhe 463
Qy      1408  GGTGCTTCTCGAAGGGGGAACATTTGATTTGTTGGAAGGCCACGAGAGAGAAATTAA 1467
Db      464  GlyAlaPheLeuLysGlyAspIleValMetPheGlnGlyAlaThrGlnGlnGlnIlyLeu 483
Qy      1468  CTGAGCGCGGAAGATGATGAATATCTAGGCTCACTTGTGCAACCGGAAATCCATATGG 1527
Db      484  LeuSerArgLysMetMetAspArgTyrlaAlaAsnPheAlaArgTrpGlyAspProAsnIly 503
Qy      1528  AACGACCTGTCTGTGGCCAGCTTATATCTGACTGACAGTACAGTACCTTCAGACTGTA 1587
Db      504  GluGlyValProLeuTrpProAlaTyrlaThrGlnSerGlnGlnTyrlaLeuLysLeuAsp 523
Qy      1588  AACATGAGCTTCGGAAGAGACTCAAAAGACCGCGGGTGAATTTTGGACCAAGCAATC 1647
Db      524  SerValSerValGlyGlnTyrlaLeuLysGlnGlnIlyValGlnPheTrpMetAsnThrIle 543

RESULT 7
Qy      095KH3  PRELIMINARY;  PRT;  361 AA.
AC  095KH3;
DT  01-DEC-2001 (TREMBLrel. 19, Created)
DT  01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT  01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE  Hypothetical protein.
OS  Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC  Cercopithecinae; Macaca.
OX  NCBI_TaxID=9541;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Temporal lobe right;
RA  Oesada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA  Suzuki Y., Sugano S., Hashimoto K.;
RL  Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC  -1- SWIMLARI: Belongs to the type-B carboxylesterase/lipase family.
CC  EMBL; AB060873; BAB46884.1; -.
DR  HSPSP; P12337; I44Y.
DR  InterPro; IPR002018; CarboxylesteraseB.
DR  Pfam; PF00135; Coesterase; 1.
KW  Hypothetical protein.
SQ  SEQUENCE 361 AA; 41032 MW; AC62AB6E78C5F1A1 CRC64;

Alignment Scores:
Pred. No.: 3,74e-116 Length: 361
Score: 1676.50 Matches: 323
Percent Similarity: 96.47% Conservative: 5
Best Local Similarity: 95.00% Mismatches: 11
Query Match: 52.89% Indels: 1
DB: 2 Gaps: 1

US-10-023-515-3 (1-1746) x 095KH3 (1-361)
Qy      724  ATACTGTCTCCATGGGCAAAAGCTTATTTCACAAAGCCATGAGAGTGGGGTGGCC 783
Db      17  IleLeuSerProMetAlaGlnGlyLeuPheHsAlaIleMetGluSerGlyValAla 36
Qy      784  ATCATCCCTTAACTGAGGCGCCATGATTAATGAGAAGAGTGAAGCACTCGAGTGTGCA 843
Db      37  IleIleProTyrlaLeuLysAlaTyrlaAspTyrlaLysSerGlnAspLeuGlnValAla 56
Qy      844  CATTTCGTGTGTAACAATGCGTCAACTCTGAGCGCTGCTGAGGTGCTGAGGACAA 903
Db      57  HisPheCyLeuLysAsnAlaSerAspSerGluAspLeuLeuArgCyLeuArgThrLys 76

```


OY 904 CCTTCAAGAGCTGCTGACCTGACGAGCAAAAGTCTTTTCACTGAGTGTGT 963
 DB 77 SeiserlysgluLeuLeuThleuSerGlnlySalysSerPheThrArgValAlaAsp 96
 OY 964 GGTGCTTCTTCTTAATGAGCTCTAGATCTATTGTCTCAGAAAGCAATTAAGCAAT 1023
 DB 97 GlyAlaPhePheProAsnGlnPheLeuGlnLeuSerGlnlyAlaPheValIle 116
 OY 1024 CTTTCATCATTCGAGGTCAATTAACGAGTGTGGCTTCTGCTGCTCATGAAGAGCT 1083
 DB 117 ProSerIleIleGlyValAlaAsnAsnIsglyGlyPheLeuLeuProMet---GlnAla 135
 OY 1084 CCTGAGATCTCTGAGTGGCTCCAAAGTCCCTTGGCTCCATCTGATTAACAAACATCCGTG 1143
 DB 136 ProGlnIleLeuSerGlySerAsnlySerleuAlaLeuIleGlnIleAsnIleLeu 155
 OY 1144 CACATCCCGCTCAGATATTGTGACCTGTGTGCTAATGATACTTCATGACAGCACTCC 1203
 DB 156 HistIleProGlnIleGlyLeuValAlaAsnGlnlyPheHlaAspIlyshIleSer 175
 OY 1204 CTGACTGAATTCGAGACAGTCTTCTGACCTTGTGAGATGTGTCTTGTGGTCCCT 1263
 DB 176 LeuThrGlnIleArgAspGlyLeuLeuAspLeuGlnlyAspValAlaPheValAlaPro 195
 OY 1264 GCACGTGATCAGAGCTCGATATTCACAGAGATGCTGTCGACCTGTACTCTCATAGTGT 1323
 DB 196 AlaLeuIleThrAlaArglyThrIleArgAspAlaGlyAlaProValIlyPheIlyIlePhe 215
 OY 1324 CGGACACCGCTCAGTGTCTTGAAGACAGCAAGCGGCTTGTGTGAAGCCGACCACTCC 1383
 DB 216 ArgHisArgProGlnIleGlyLeuAspIlySerPheIleValAlaAspIleAla 235
 OY 1384 GATGAAGTCCGCTTGTGTCTGCTGCTGCTTCTGAAAGGGGACATTTATGTTGAA 1443
 DB 236 AspGlnValAlaArgPheValAlaPheGlyGlyAlaPheLeuIlyGlyAlaIlePheGln 255
 OY 1444 GGAGCGACGAGGAGGAGGAGGAGTACTGAGCGGAGAGATGAATTAATCTGGCTCTT 1503
 DB 256 GlyAlaThrGlnIleGlnIleGlnIleLeuLeuSerArgIlySerMetLeuIlyPheIleThrPhe 275
 OY 1504 GCTCGAACCAGGAAATCTTAATGGAAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1563
 DB 276 AlaArgThrGlyAlaAsnProAsnGlyAspAspLeuProLeuIlePheIleThrPheIleThr 295
 OY 1564 GAGCAGTACCTTCAGCTGAGTGAATGATGAGCTTGAAGCTGAGAGCTGAGAGCTGAG 1623
 DB 296 GlnGlnIlyIleGlnIleLeuAspLeuAsnMetSerIleGlnArgLeuIlyGlnProArg 315
 OY 1624 GTGATTTTGGACACAGACCATCCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
 DB 316 ValGlnPheIlePheIlePheIlePheIlePheIlePheIlePheIlePheIlePheIle 335
 OY 1684 CCTCTTCTCTCTTCAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1743
 DB 336 ProLeuSerPheLeuIlePheLeuSerLeuLeuGlnIleProPhePhePheSerGlyValPro 355
 RESULT 8
 SASB ANAPL STANDARD: PRT: 557 AA.
 AC 004751:
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DB Fatty acyl-CoA hydrolase precursor, medium chain (EC 3.1.2.-)
 DE (Thioesterase B).
 OS Anas platyrhynchos (Domestic duck).
 OC Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.
 OX NCBI_TaxID=8839;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-65.
 RC TISSUE=urological gland;

RX MEDLINE=93300823; PubMed=8314791;
 RA Hwang C.-S., Kolatukudy P.E.;
 RT "Molecular cloning and sequencing of thioesterase B cDNA and
 RT stimulation of expression of the thioesterase B gene associated with
 RT RT hormonal induction of peroxisome proliferation.";
 RL J. Biol. Chem. 268:14278-14284(1993).
 CC -1- FUNCTION: Fatty acid biosynthesis chain termination and release of
 CC the free fatty acid product is achieved by hydrolysis of the thio
 CC ester by a thioesterase. This thioesterase may be associated with
 CC peroxisome proliferation and may play a role in the production of
 CC 3-hydroxy fatty acid diester pheromones.
 CC -1- TISSUE SPECIFICITY: Highest levels in uropygial gland, much lower
 CC in liver and kidney.
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL: L05493; AAA49223.1; -.
 DR PIR: A47162; A47162.
 DR HSSP: O7540; 1K4Y.
 DR InterPro: IPR002018; Carboxylesterase.
 DR InterPro: IPR003579; Ser_estr.
 DR Pfam: PF00135; Coesterase; 1.
 DR PROSITE: PS00122; CARBOXYLESTERASE B_1; 1.
 DR PROSITE: PS00941; CARBOXYLESTERASE B_2; 1.
 KW Direct protein sequencing; Fatty acid biosynthesis; Hydrolyase; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 557 Fatty acyl-CoA hydrolase.
 FT ACT_SITE 227 227 Acyl-ester intermediate (By similarity).
 FT ACT_SITE 345 345 Charge relay system (By similarity).
 FT ACT_SITE 460 460 Charge relay system (By similarity).
 FT DISULFID 93 122 By similarity.
 FT CARBOHYD 476 476 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 557 AA; 61637 MW; 03B35D90A037F680 CRC64;
 Alignment Scores:
 Pred. No.: 6,45e-88 Length: 557
 Score: 1298.50 Matches: 265
 Percent Similarity: 62.77% Conservative: 79
 Best Local Similarity: 48.36% Mismatches: 177
 Query Match: 40.96% Indels: 27
 DB: 1 Gaps: 8
 US-10-023-515-3 (1-1746) x SASB_ANAPL (1-557)
 OY 88 ACTGGGCTTCTGCTGAAAGGCGACAGAGAACACAGGCTGGATGATTCAGGGCAAG 147
 DB 23 ThrGlyGlnlyAlaGlnIleGlnIleProGlnIleValIleThrAsnIlyGlySerValArgIlyIlyr 42
 OY 148 CAAGTCACGTGCTGCGAAGCCCTGCTGGAACGTGCTTCTGAGAGTCCCTTGTGT 207
 DB 43 GlnValIlyValAlaAsnAlaIleGlnIlyArgSerValAlaValIleLeuGlyLeuPhePheAla 62
 OY 208 GCTCCCGCTGAGATCCCTGATTTTACGAACCGGAGCTGATCCCTGAGATAAC 267
 DB 63 LysProProValGlyProLeuArgPheSerGlnIleProGlnIleProGlnIleProIlyPheGly 82
 OY 268 TTGCAGAGCCACCTCTTACCTTAATTTGCTGCTCCAGAACTCAGAGTGTGCTCTTA 327
 DB 83 ValArgAspAlaAlaSerIlyrProPheMetCysLeuGlnIlePheValIleValIleValIle 99
 OY 328 GATCAACACATGCTC-----AAGTGCAATTACCCGAATTCGAGATGCA 372
 DB 100 GlnGlnIlyIleSerAspAlaIleThrAsnArgIlyGlnIlyValAlaArgLeuGlnIleSer 119
 OY 373 GAAGACTGCTCTACCTGATCAATCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432

D	b		120	GluAerpyrValenUryleuaenValTyrThrProValSerThreGlutGlnGlyuBlen	135
O	y		433	CCGGTCCTTGCGTGGTTCCAGAGAGTGCCTTCAACGTGGCTCAGCTCCATCTTGAAT	492
D	b		140	ProValPheValTrpIleHseGlyGlyLeuValSerGlyAlaIAserSerTyraar	159
O	y		493	GGGCGCGCGCTGGGTGCCTTAGAGAAAGTGTGGTGGTGGCGCATGCCGATAGA	552
D	b		160	GlySerAlaLeuAlaIAPheAerAenValValThrIleGlnTyraGlyeGly	179
O	y		553	ATAATTGGTTCTTCAACACATGGATCAGATCTCCGGGAACTGGGCTCTCAAGAC	612
D	b		180	IleAlaGlyTyrPheSerThrGlyAerPyHAlaArgGlyAenTrpGlyTytleuar	199
O	y		613	CAGTGGCTGTCTGTCCCTGGGTCGAGAAACAATGAGATCTTCGGATGGGACCCGAC	672
D	b		200	GlnValAlaIaleuGlnTrpIleGlnGlnAmlerIleIAPheArGlyAerProGly	219
O	y		673	TCTGTGACCATTCTTGCGAGTCCGGGGAGCATAGTGTCTTAAGTCTTAATGCTCT	732
D	b		220	SerValThrIlePheGlyGlnSerIleaglyGlyValSerValSerAlaLeuValIeuer	239
O	y		733	CCCATGGCCAAAGGCTTATTCCAAAGCATATGAGATGGGGTGGCCATC--ATC	789
D	b		240	ProLeuAlaIaGlyLeuPheHsrValaIleSerGlnSerGlyThrAlaValArgIle	259
O	y		790	CCTTACCTGAGGCCCATGATTATGAGAAGAGTGAAGCATGAGGTGGTGGACATTC	849
D	b		260	LeuPheThrGlu-----GlnProGlnGlnAlaIeGlnArgIleAlaIala	275
O	y		850	TGTGTGAACANtCGTCAACTGTGAGCCCTGCTGAGTGCCTGAGCAAAAACCTCC	909
D	b		276	AlaGlyCyseGlnUlysSerSerSerAlaIaleuValGlcYleuArgGlyuIbserThru	295
O	y		910	AAGGAGCTG----CTGACCCCTCACGCCAGAAAAACAATCTTCACTGAGATGGTAT	963
D	b		296	AlaGlnMeGlnGlnIleThrLeuLysMetProPrometPheIleSerAlaSerLeuar	315
O	y		964	GGTGGTTTTCTTCTTAATGAGCCTCTAGATCTATGTTCTCAGAAAGCATTTAAACAAT	1023
D	b		316	GlyValPhePheProLySerProArgGlnLeuIeuerSerGlnUyValIleAmAlaVal	335
O	y		1024	CCTTCCATATGAGGATCAATAACACAGATGGTCTTCTGCTGCT--ATGAAGAG	1088
D	b		336	ProTyrlleIleGlyAlaAmnAnCyGlnPheGlyTrpIleuProArgMetCelys	355
O	y		1081	GCTCTGAGATCCTCAGTGCCTCAACAGTCCCTGTGCCTCCATCTGATCAAAACATC	1140
D	b		356	PheProGlnPheThrGlnUylLeuGlnUyAerValAlaArgGlnValIeueGlnSerThr	375
O	y		1141	CTGCACATC-----CCGCTCAAGATTTGCACTCTGTGGCTAATGAATAC	1185
D	b		376	LeuAlaIeuerSerPheUyGlyAlaProSerAerIleValAerUyValTyrAmGlnUyX	395
O	y		1186	TTCCATGACAAAGACCTCCCTGACTGAATTCGAAACAGTCTTGGAGCTGTGGAGAT	1245
D	b		396	IleGlyValAlaGlnAenArgAlaGlnValAlaArgPheGlyLeuIeuaAerSerIleAlaAer	415
O	y		1246	GTTGTTCTTTGGTCCCTGCACTGATCAACGCTGATATCACAAGATCTGTGCACT	1309
D	b		416	ProLeuPheValPheSerAlaValGlnValAlaArgHIsnIlaArgAerAlaGlyAenPro	435
O	y		1306	GTTATCTTCTAAGATTTCCGAGCGGCTCAGTCCCTTGAAGAACAAGACCGGCTTTT	1365
D	b		436	ValTyrPheTyrglnPheGlnHIsnArgProSerSerAlaIaGlyValaIProGlnHe	455
O	y		1366	GTCMAAGCCGACACGCTGATGAAGTCCGTTTGTTCGGGTGAGCTTCTTGAAGGG	1425
D	b		456	ValUyValAerHIsnAlaAerGlnUylleAlaPheValPheGlyUyAerPheIeuaIaGly	475
O	y		1426	GACATTTGTTATGTGAAAGAGCCACGAGAGAGAAAGTTACTGACCCGAAATGATG	1485
D	b		476	Aen-----AlaThrGlnGlnGlnUylalUyBlenSerAerThyAerMet	489

Qy	1466	AAATATCTGGGCTACTCTTTGGCTCGAACC	CGGGAAATCTTAATGGAAACGACCTCTCTGTGG	1545
Db	430	LYSTYTRTPHRAENPHEALALAYASbNLYASbNP	ProAbnNglyIugJlyLeuValH1stTP	509
Qy	1546	CCAGCTTAATCTGACTGACGACATACCTCCAG	CTGACCTGAATCTGAACATGAGCCTCGACAG	1605
Db	510	ProGlnTyrAPMeLaepGluArgTyrLeuGlnL	LeaPLeuThrGlnLysAlaAlaLys	529
Qy	1606	AGACTCAAAAGAACCGCGGGTGAATTTTGGAC	CGACACCATCCCTCGATCTGTCTGCC	1665
		:::	:::	
Db	530	LYSLLeuYbGluArgLysMetGluPheThrMet	GlnLeuThrGlnLysIleMetSerAsp	549
Qy	1666	TCGCACATGCTCCACAGTCTCTT	1689	
Db	550	ArgArgArgLysH1stH1stH1stH1stH1st	557	
RESULT 9				
ID	Q91WG0	PRELIMINARY;	PRT;	561 AA.
AC	Q91WG0			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)			
DE	Carboxylesterase 2 (Carboxylesterase M3 precursor).			
GN	Name=Ces2; Synonyms=ces2A3;			
OS	Mus musculus (Mouse);			
OC	Eukaryota; Metazoa; Chordata; Cranicia; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	NCBI_TaxID=10090;			
	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N. TISSUE=Colon, and Kidney;			
RX	MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haefl F.,			
RA	Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Ueda T.B., Tohilyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Falley J., Helton E., Kettelman W., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madsen A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Kryzanski M.I., Skalka U., Smilins D.E., Schnerch A., Scheln J.E.,			
RA	Jones S.J., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N. TISSUE=Kidney;			
RA	Strausberg R.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N. TISSUE=Colon;			
RA	Strausberg R.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N. TISSUE=Colon;			
RA	Strausberg R.;			
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FVB/N. TISSUE=Colon, and Kidney;			
RA	Strausberg R.;			

QY	211	CCCCCGCGGGATCCCTGGCATTTTAACGACCCGACGCTGCATCGCCCTGGGATTACTTG	270
Db	66	ProProValGlyProLeuAlaPheLeuAProProGluAlaProGluProTrpSerGlyVal	85
QY	271	CGAAGACGACCTCCCTCAATTTGTGGCTCCAGACTCAGAGTGAGCTG---CTCTTA	327
Db	86	ArgArgSerGlyThrSerHisProHisMetCysLeuGlnIleAsnAraPheMetLeuMetGlySerGlu	105
QY	328	GATCAACACATGCTCCAAAGTGCATTTACCCGAAATTCGAGTGTCCAGAAAGACTGCTCTAC	387
Db	106	AspLeuIysMetLeuMetAsnLeuIleLeuProProIleSerMetSerGluAspCysLeuTrp	125
QY	388	CTGAACATCTATGGCGCTGCCACGCGGCATACAGGCTCCAAAGCTCCCGCTTGTGGTGG	447
Db	126	LeuAsnIleTrpValProIleHisIleAsnIleGluGlySerAsnLeuProValMetValTrp	145
QY	448	TTCCGAGAGGTGCCTTCAGACTGCTCCAGCTGCATCTTTGATGGTCCGCGCTGACT	507
Db	146	IleHisGlyGlyAlaLeuThrValGlyMetAlaSerMetTrpAspGlySerMetLeuAla	165
QY	508	GCTATGAGACGCTGCTGGTGTGGTCCGTCAGATACCGGCTAGAGAAATTTGTGTTCTTC	567
Db	166	AlaThrGluAspValValValAlaIleGlnTrpArgLeuGlyValLeuGlyPhePhe	185
QY	568	ACCAATGGGATCAGACTGCTCCGGGAACTGGCGCTTCAGAGACCAAGTGGACTGCTGG	627
Db	186	SerThrGlyAspGlnHisAlaIleAspGlyAsnProAspArgValThrIlePhe	205
QY	628	TTCCTGGCTCCAGAAAGATCGAGTTCCTCGTGGGGACCCAGCTGTGATGACATCTTT	687
Db	206	ArgTrpValGlnGlnAsnIleValHisPheGlyGlyAsnProAspArgValThrIlePhe	225
QY	688	GGCAGAGTCCGGGGAGCCATAAGTGTCTTAGTCTTAATGCTGCCCATGGCCAAAGC	747
Db	226	GlyIleSerAlaGlyGlyThrSerValSerSerHisValValSerProMetSerGlnGly	245
QY	748	TTATTTCCAAAGCCATCATGGAGAGTGGGGTGGCCATCATCCCTTACCTGAGGCCCAT	807
Db	246	LeuPheHisGlyAlaIleMetGlySerGlyValAlaValLeuProAspLeuIleSerSer	265
QY	808	GATTATGAGAAAGTGAAGACTCTGAGAGTGGTGGACATTTCTGTGTTAACAAATGCCCA	867
Db	266	SerSerGlyMetValHis-----ArgIleValAlaAsnLeuSerGlyCysAlaAlaVal	283
QY	868	GACTCTGAGGCTCTGAGTGCCTGAGAGCAAAACCTCCAAAGAGCTGCTGCTC	927
Db	284	AsnSerGlyThrLeuMetCysCysLeuAspGlyValAsnGlnAlaGluMetLeuAlaIle	303
QY	928	AGCCAGAAACAAAGTCTTTCACTCGAGTGGTGTGATGGTCTTTCTTAATGAGCT	987
Db	304	AsnIysValPheIysIleIleProGlyValValAspGlyGlnPheLeuProIleHisPro	323
QY	988	CTAGATCTATGTCTCAGAAAGCATTTAAAGCATTCCTTCATCATGTGAGTCAATAC	1047
Db	324	GlnGluLeuMetAlaSerTrpAspPheHisProValProSerIleIleGlyIleAsnAsn	343
QY	1048	CACGAGTGGCTCTCCGCTGCT-----	1071
Db	344	AspGlyTrpGlyTrpIleLeuProThrIleMetAspProAlaGlnIysIleGluGluIle	363
QY	1072	---ATGAAGAGGCGCTCGAGATCTCTCAGTGGCTCCAAAGTCCGCTTGCCTCATCTG	1128
Db	364	ThrArgGlySerThrLeuProAlaValLeu-----LysSerThrAlaLeuIysMet	379
QY	1129	ATACAAACATCTGACATCCCGCTCAGATTTTGACCTTTGGCTAATGATATCTTC	1188
Db	380	Met-----LeuProProGluCysGlyAspLeuLeuMetGluGlyTrpMet	394
QY	1189	CATGACAAAGCACTCCCTGAATGAAATCCGAGACACTCTTTCGACTTGGCTGGAGATGTG	1248
Db	395	GlyAspThrGluAspProGluThrLeuGlnAlaGlnPheArgGlnMetLeuIleAspPhe	414

Oy		1249	TTCCTTGTGGCCCGCTGACAGATGCAGAGCTCAGATGCAGAGTAGTGCGGTGCACCTGTC	13068
Dd		415	MethveValIleProAlaLeuGIlnValAlaHisPheGlnTrg--SerHISAlaProval	433
Oy		1309	TACTTCATGAGTTTCCGCCACCAGCCGCTCAGTGCCTTTGAAGAACAAGAACCGGCTTTTGTCT	13668
Dd		434	TyrThrTyrglunPheGlnHisArgProSerPheHeLysAsnRheAsgPrProTy-Val	453
Oy		1369	AAAGCCCAACAAGCTGATGAAGTCGCTTGTGTCTGCGTGGCTTCCTGAAGGGGAC	1428
Dd		454	LysAlaAsnHisGlyAsnArgJulilePheLeuValPheGlyTyrglnPhe-----GlyAsn	471
Oy		1429	ATTGTTATGTTCCGAAGAGCAACGAGAGAGAGAAAGTTACCTGAGCCGGAAGATGATGAA	1488
Dd		472	IleLysLeu-----ProTyrrThglunglInglunGlunLeuSerArgArgIleMetLys	489
Oy		1489	TACGAGGCTACCTTGTGCTGCAACCGGAGATCTTAATGGGAACGACTGCTCTGTGGCCA	1548
Dd		490	TyrrThrAlaAsnRheAlaArgHisGlyAsnProAsnSerGluGlyLeuProTyrrPro	509
Oy		1549	GCTTATATCTGACATGACAGCATCTTCCAGCTGGACTTGAACATGACCTTCGACAGAGA	1608
Dd		510	ValMetAsnHisArgGluGlnTyrlenGlnLeuAsnPrIleGlnProSerValGlyArgAla	529
Oy		1609	CTCAAAGAACCGCGGGTGGATTTTGGACCAAGCACCATCCCC	1650
Dd		530	LeuLysAlaArgLeuGlnPheTrpThrLysThrLeuPro	543
 RESULT 11 035533 PRELIMINARY; PRT; 559 AA.				
ID	035533	AC	035533;	
DT	01-JAN-1998	(TREMBLrel_05, Created)		
DT	01-JAN-1998	(TREMBLrel_05, Last sequence update)		
DE	01-MAR-2004	(TREMBLrel_26, Last annotation update)		
DE	Carboxylesterase precursor (EC 3.1.1.1).			
OS	Mesocricetus auratus (Golden hamster).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
OC	Mesocricetus.			
OX	NCBI_TaxID=10036;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Syrian golden; TISSUE=Liver;			
RA	Sone T., Wang C.Y.;			
RT	"Microsomal amidases and carboxylesterases."			
RL	(In) Guengerich F.P. (eds.);			
RL	COMPREHENSIVE TOXICOLOGY VOLUME 3, Bioc transformation, pp.265-281,			
RL	Pergamon, Oxford (1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Syrian golden; TISSUE=Liver;			
RA	Sone T., Ishida Y., Takabatake E., Wang C., Isobe M.;			
RL	Submitted (May-1995) to the EMBL/GenBank/DBJ databases.			
CC	-1-SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.			
DR	EMBL; D50577; BA23605.1; -.			
DR	HSSP; P12337; IK4Y.			
DR	GO; GO:0004091; F:carboxylesterase activity; IEA.			
DR	GO; GO:0016787; F:hydrolase activity; IEA.			
DR	InterPro; IPR002018; Carbesteraeb.			
DR	InterPro; IPR000886; ER_target_S.			
DR	InterPro; IPR00379; Ser_eatrb.			
DR	PROSITE; PS00123; CARBOXYLESTERASE_B_1; 1.			
DR	PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.			
DM	PROSITE; PS00014; ER_TARGET; UNKNOWN_1.			
KW	Hydrolase; Signal.			
FT	SIGNAL	1	27	Potential.
FT	CHAIN	28	559	carboxylesterase.
SQ	SEQUENCE	559 AA;	62126 MM;	SB68D4FPA47C6AD8 CRC64;

Alignment Scores:

Pred. No.:	2.82e-86	Length:	559
Score:	1276.50	Matches:	262

Percent Similarity: 64.58% Conservative: 79
Best Local Similarity: 49.62% Mismatches: 172
Query Match: 40.27% Indels: 15
DB: 2 Gaps: 9
US-10-023-515-3 (1-1746) x 035533 (1-559)

82 GAAAAATCGGGCTTGTGCTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAG 141
Db GlycinapserProser-----ProtleghantthiethrGlycinalArg 42
26 GlycinapserProser-----ProtleghantthiethrGlycinalArg 42
142 GGCAAGCAAGTCAGTCTGTGGAAAGCCCTGTGCTGAAGCGTTCCTCGAGTCCC 201
Db GlythgleValAlaValLysAspThrGluValAspValAlaValLysPheLeuclYlePro 62
43 GlythgleValAlaValLysAspThrGluValAspValAlaValLysPheLeuclYlePro 62
202 TTGTGCTGCTCCCGCTGGATCCTGGATTAAGAACCCGACGCTGACGCTGG 261
Db PheAlaValProValAlaValProLeuAlaValProValAlaValProValAlaValPro 82
63 PheAlaValProValAlaValProLeuAlaValProValAlaValProValAlaValPro 82
262 GATAACTGGAGAGCCACCTACCTAATTTGTGCTCCAGAACTCAGAGTGGCTG 321
Db SerGlyValAlaValAspGlyValAspSerPheProAlaMetCysLeuGlnAspAspMetMet 102
83 SerGlyValAlaValAspGlyValAspSerPheProAlaMetCysLeuGlnAspAspMetMet 102
322 CTCTTAAT---CAACACATGCTCAAGGTCAATTAACCGAAATTCGAGTGCAGAGAC 378
Db AsnSerGluGlyLeuLysMetCileuLysLeuLysMetCileuLysLeuLysMetCileuLys 122
103 AsnSerGluGlyLeuLysMetCileuLysLeuLysMetCileuLysLeuLysMetCileuLys 122
379 TGGCTTACCTGCAAGATCATGCGCTGCCACCGCGCATACAGGCTCCAGTCCCGCTC 438
Db CysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuVal 142
123 CysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuVal 142
439 TTGTGCTGCTCCCGCTGGATCCTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498
Db MetValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 162
143 MetValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 162
499 GCGCT 558
Db MetLeuValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 182
163 MetLeuValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 182
559 GGTTCCTTACACATGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 618
Db GlyPhePheSerThrGlyAspGlnValAlaValLysLeuValLysLeuValLysLeuValLys 202
183 GlyPhePheSerThrGlyAspGlnValAlaValLysLeuValLysLeuValLysLeuValLys 202
619 GCT 678
Db AlaAlaLeuValGlyValGlnGlnAlaValLysLeuValLysLeuValLysLeuValLysLeuVal 222
203 AlaAlaLeuValGlyValGlnGlnAlaValLysLeuValLysLeuValLysLeuValLysLeuVal 222
679 ACCATCTTGGCGAGTCCGCGAGCCATAAGTGTTCATCTTAATCTGCTCCATG 738
Db ThrLysPheGlyGlnSerAlaGlyLysSerValSerSerThrValValSerProMet 242
223 ThrLysPheGlyGlnSerAlaGlyLysSerValSerSerThrValValSerProMet 242
739 GCCAAGGCTTATCCCAAGACCATCATGAGAGTGGGGTGGCGCATGCTCCATGCTG 798
Db SerLysGlyLeuPheLysGlyAlaLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 262
243 SerLysGlyLeuPheLysGlyAlaLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 262
799 GAGGCCCATGATTTGAGAGAGTGAAGGACCTGAGAGTGTGTCACATTTTGTGTGTAAC 858
Db LysSerSerSerSerGlyMetVal-----TyrThrLysValAlaAlaAlaLysSerAspPys 280
263 LysSerSerSerSerGlyMetVal-----TyrThrLysValAlaAlaAlaLysSerAspPys 280
859 AATGCGTCAGACTGTAGGCTGTGCTGAGTGGCTGAGAGCAAAACCTCCAGAGAGCTG 918
Db AlaAlaValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 300
281 AlaAlaValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 300
919 CTGACCGCTGAGCGAGAAACAAAGCTTTCACCTGAGAGTGTGTAAGTGTGCTTCTTCT 978
Db LeuAlaValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 320
301 LeuAlaValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 320
979 AATGAGCTCTAGATCTATTGTCTCAGAAACATTTAAAGCAATTCCTTCATCATCGGA 1038
Db LysAlaPheProGlnGlnLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 340
321 LysAlaPheProGlnGlnLysLeuValLysLeuValLysLeuValLysLeuValLysLeuValLys 340
1039 GTCAATATACCAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
Db GTCAATATACCAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092

Db ValAlaAsnAspArgLysTyrGlyTyrLeuValProValLysMetGlySerAlaGlnGluLe 360
341 ValAlaAsnAspArgLysTyrGlyTyrLeuValProValLysMetGlySerAlaGlnGluLe 360
1093 CTCAGTGGCTCCACAGAGTCCCTT-----GCCCTCATCTGATACAAACATCCTGCAC 1146
Db LysGluLysLeuValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuVal 380
361 LysGluLysLeuValAlaValLysLeuValLysLeuValLysLeuValLysLeuValLysLeuVal 380
1147 ATCCCGCTCAGATTTATTCGAGCTTGTGCTAATGATCTTCATACCAAGACCTCCCTG 1206
Db LysProProGluLysGlyValAspLeuLysMetGluLysTyrMetGlyValAspGlnLysPro 400
381 LysProProGluLysGlyValAspLeuLysMetGluLysTyrMetGlyValAspGlnLysPro 400
1207 ACTGAATCCGAGACAGTCTTCTGAGCTGCTTGGAGATGTGTTCTTGTGCTCCGCA 1266
Db GlnThrLeuGlnAlaGlnPheArgGluMetMetGlyAspPheThrPheValLysProSer 420
401 GlnThrLeuGlnAlaGlnPheArgGluMetMetGlyAspPheThrPheValLysProSer 420
1267 CTGATCAGAGCTGCATATCAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
Db LeuGlnValAlaValLysPheGlnArg---AlaValAlaProValTyrPheTyrGluPheGln 439
421 LeuGlnValAlaValLysPheGlnArg---AlaValAlaProValTyrPheTyrGluPheGln 439
1327 CACCGGCTCAGTCTTGTGAAGACAGAGGCGCTTGTGCAAGCCGACACGCTGAT 1386
Db HisArgProSerPhePheLysAspThrLysProProHisValLysValAspGlnLysPro 459
440 HisArgProSerPhePheLysAspThrLysProProHisValLysValAspGlnLysPro 459
1387 GAGTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
Db GluPhePhePheValLysGlyValAsn---LeuLysCysGly-----LysLysPhe---Ala 475
460 GluPhePhePheValLysGlyValAsn---LeuLysCysGly-----LysLysPhe---Ala 475
1447 GCCACGAGAGAGAGATTAATGAGCGGAGATGAGAAATGAGTGGCTGCTGCTGCTGCTGCTGCT 1506
Db SerThrGluGluGluGluLysLeuLysSerArgLysMetCysLysTyrPheValAspPheAla 495
476 SerThrGluGluGluGluLysLeuLysSerArgLysMetCysLysTyrPheValAspPheAla 495
1507 GCAACCGGAGATCTTAATGAGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
Db ArgHisGlyAsnProAsnSerGluAsnLeuProTyrLysProMetLysAspPheAlaGln 515
496 ArgHisGlyAsnProAsnSerGluAsnLeuProTyrLysProMetLysAspPheAlaGln 515
1567 CAGTACTCCAGCTGAGCTTGAATGAGTGCCTGCAAGAGATCAGAAACCGCGCTG 1626
Db GlnTyrLeuGlnLeuAspLysLysProAlaValGlyValArgLysLeuValAlaArgArgLeu 535
516 GlnTyrLeuGlnLeuAspLysLysProAlaValGlyValArgLysLeuValAlaArgArgLeu 535
1627 GATTTTGGACACGACCATCCG 1650
Db GlnPheThrLysLysLeuPro 543
536 GlnPheThrLysLysLeuPro 543

RESULT 12
Q8BM97 PRELIMINARY; PRT; 559 AA.
AC Q8BM97;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus adult male colon cDNA, RIKEN full-length enriched
DE library, clone:903624J02 product:similar to CARBOXYESTERASE (EC
DE 3.1.1.1) (A11-ESTERASE) (B-ESTERASE) (MONOBUTYRASE) (COCAINE ESTERASE)
DE (PROCAINE ESTERASE) (METHYLBUTYRASE).
GN Name=903624J02Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RP STAIN=C57BL/6J; TISSUE=Colon;
RC MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RX Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RN Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RP STAIN=C57BL/6J; TISSUE=Colon;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RX RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RL "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
[41]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
[51]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Atzawa K., Nagaoke S., Saeki N., Carninci P.,
RA Kono H., Akiyama Y., Nishi K., Kikunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiwa K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara B., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
[61]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Colon;
RA Adachi U., Atzawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hisokane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kikawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Muraoka M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sato H.,
RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AK033563; BAC28361.1; -.
DR HSSP; P12337; 1K4Y.
DR MGD; MGI:2443170; 9030624L02R1K.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR002018; Carboxylesterase.
DR InterPro; IPR003719; Ser esterase.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase.
KW PROSITE: PS00941; CARBOXYLESTERASE_B_2; 1.
SQ SEQUENCE 559 AA; 62305 MW; 8C54D21522C252C4 CRC64;

Alignment Scores:
Pred. No.: 7,24e-86 Length: 559
Score: 1271.00 Matches: 260
Percent Similarity: 63.30% Conservative: 78
Best Local Similarity: 48.69% Mismatches: 166
Query Match: 40.09% Indels: 30
DB: 2 Gaps: 8

US-10-023-515-3 (1-1746) x Q8BM97 (1-559)
QY 91 GGGCTTCTGCTGAAGGGCCAGAGAACACAGAGCTGGAGTTCAGAGGCAAGCA 150
DB 26 G1GTAAGSerAlaSerProIleArgValThrHisGlnValAlaGlySerLeu 45
QY 151 GTGACGTGCTGGAGAGCCCTGTGCTGTAACGTGCTTCCTGGAGTCCCTTGTGTCT 210

DB 46 ValHisValIysAspThrAspIleAlaValHisThrPheLeuGlyIleProPheAlaIys 65
QY 211 CCCCCTGGAGATCCCTCCGATTAACGAACCCGAGCCCTGCATCCCTGGATTAAC 270
DB 66 ProProValGlyProIleuAlaPheAlaProProGluAlaProGluProIleSerGlyVal 85
QY 271 CGAAGAGCCACCTCTTACCTTAATTGTGCTCCGAGACTCGAGATGAGTGTGCTG 327
DB 86 ArgAspGlyThrSerHisProAspMetCysLeuGlnAlaAspAspLeuMetCysIleGly 105
QY 328 GATCAACATGCTCCAGTGCATTAATCCGAAATTCGGAGTGTCAAGAGATGCTCTTAC 387
DB 106 AspLeuIysMetLeuAlaLeuIleLeuProProIleSerMetSerGlyAspCysLeuIyr 125
QY 388 CTGAACATCTATGCGCTGCCCGCCAGCCATACAGGCTCCAGACTCCCTGTGCTGTG 447
DB 126 LeuAsnIleIyrValProAlaHisAlaHisGlnGlySerAsnLeuProValMetValTyr 145
QY 448 TTCCAGAGAGGCTTCGAAGCTGCTCAGCTCCATCTTGTGATGGTCCGCTGTGCT 507
DB 146 IleHisGlyGlyAlaLeuThrValGlyMetAlaSerMetIyrAspGlySerMetLeuAla 165
QY 508 GCGTATGAGAGCTGCTGCTGTGCTGCTCCAGTACCGGCTTACGAAATTTGGTTCTTC 567
DB 166 AlaThrGluAspValValValAlaAlaIleGlnIyrArgLeuGlyValLeuGlyPhePhe 185
QY 568 ACCACATGGAGATCAGATGCTCCGCGGAACTGGGCTTCGACAGACACAGTGTGCTGTG 627
DB 186 SerThrGlyAspGlnHisAlaIleAlaIleGlyValSerIyrLeuAspGlnValAlaIleLeu 205
QY 628 TCCGTGCTCCAGAAAGATCAGATGCTTCCGTGGGAGCCCGAGCTCTGTGACATCTTT 687
DB 206 ArgTyrValGlnGlnAlaHisIleValHisPheGlyValAspProAspArgValThrIlePhe 225
QY 688 GCGGAGTCCGCGGAGCCATAGTGTCTTCTAGTCTTATGATGCTCCATGCGCAAGGC 747
DB 226 GlyIleSerAlaGlyGlyThrSerValSerSerHisValValSerProMetSerGlnGly 245
QY 748 TTATTCACAAAGCCATGAGAGAGTGGGCTGCGCATATCCCTTACCTGGAGCCCAT 807
DB 246 LeuPheHisGlyAlaIleIleMetGluSerIleValAlaValIleProAspLeuIleSer 265
QY 808 GATTATGAGAGAGTGAAGACCTGACAGTGGTGTGACATTTCTGTGTAACAATGCTGCA 867
DB 266 SerSerGluMetValHis-----ArgIleValAlaAsnLeuSerGlyCysAlaIleVal 283
QY 868 GACTGTGAGCCCTGCTGAGGTGCTGAGAGCAAAACCTCCAGAGACTGCTGACCTCT 927
DB 284 AsnSerGlyThrIleuMetCysCysLeuArgGlyValAsnGlnAlaGluMetLeuAlaIle 303
QY 928 AGCCAGAAACAAAGCTTTTCACTGAGTGGTGTGATGCTTTCTTCTTATAGAGCT 987
DB 304 AsnIysValPheIysIleIleProGlyValValAspGlyGluPheLeuProIyHisIyrPro 323
QY 988 CTAGATCTATGTCTCAGAAAGCATTAAGCAATTCCTTCATGATGAGTCAATAAC 1047
DB 324 GlnGluLeuMetAlaSerIysAspPheHisProValProSerIleIleGlyIleAsnAsn 343
QY 1048 CACGAGTGTGCTTCTGCTGCTCT----- 1071
DB 344 AspGlyIyrGlyTyrIlePheLeuProThrIleMetAspProAlaGlnIleIleGluIle 363
QY 1072 ---ATGAAGAGGCTCTGAGATCTCAGTGGCTCAAAAGTCCCTGCTTCATCTG 1128
DB 364 ThrArgIysThrIleuProAlaValLeu-----LysSerThrAlaLeuIysMet 379
QY 1129 ATACAAACATCTTCAGATCCCGCTCAGTATTTGCACTTGTGGCTAATGAATATCTTC 1188
DB 380 Met-----LeuProGluIyrGlyValAspLeuMetGluGlnIyrMet 394
QY 1189 CATGACAAGCACTCCCTGATCGAATCCGAGACAGTCTTCTGAGCTTCTGAGATGTG 1248

Db 395 Glyaerphrthgluaerproglutthireuetaiaaglnphehargglumellyaerph 414

Qy 1249 TTTCTTGTGTCCTGCACTGATCAACGCTGCATATCAACAGATGCTGTGCACCTGTC 1308

Db 415 Metcpehvalileproalaileuglnvalalanhiserghnarg---Serh;alaproval 433

Qy 1309 TACTTCTATGAGTTTCGGACCCGGCTCAGTGCCTTTGAAGACAGAACCCGGCTTTGTC 1368

Db 434 Tyrrheurygluupheglnhlayrproserphehellyaerphneatgprotyrrhval 453

Qy 1369 AAAGCCGACACGCTGATGAATCCGCTTGTTCGGTGTGCTCTTGCATGAAGGGGAC 1428

Db 454 Lysalaaerphleaglyaerphlulepheuvalphegilytyrglnphe-----Glyasn 471

Qy 1429 ATTGTATGTTTCSAAGAGACCCAGAGAGAGAAATTACTGAGCCGGAATGATGATAA 1488

Db 472 Ilelybleu-----ProtyrthrsluglnluglnleuSerayargmeteclys 489

Qy 1489 TACTGGGATCACTTGTGCTCGAACCGGGAAATCTTAATGGAAAGCATCTGTCTGTGGCCA 1548

Db 490 TyrrargalaaenphelalarghileglyasnProhansergluglyuenProtyrtrpPro 509

Qy 1549 GCTTATATATGCACTGAGACAGTACCTCCAGCTGGAATTGAACATGAGCCTCGACAGAA 1608

Db 510 ValmetcpehphlaaergluglnlytyleuglnleuSerhileglnproservalglyargla 529

Qy 1609 CTCAAGAACCCGGGTGGATTTTGGACCGACGACCATCCCC 1650

Db 530 LeullysalaargargyleuglnphehtrpThryshyrthleupro 543

```

Oy 775 GGGGTGGCCATCATCCCTTACTGAGAGGCCCATGATTATGAGAGAGTGAAGAGCTGGAG 834
    |||||.....|
Db 252 GYValAlaLeuLeuProTyLeuLeuIerhAerhPheSerGluMetValSer-----Thr 269
Oy 835 GTGGTTGACATTTCTGTGTGTAACAATGCGTCAGACTTGAGGCGCTGTGAGTGGTCTG 894
    |||||.....|
Db 270 ThrValAlaTyLeuSerGlyCySerGluMetAerSerGluAlaLeuValArgCySer 289
Oy 895 AGGAGAAAACCTCCAGAGAGCTGTGAGACCCCTCACCCAGAAAACAAGCTTTCATCTGA 954
    |||||.....|
Db 290 ArgGlyTySerGluAlaGluLeuAlaIleAsnTyLeuValGlnMetIleProAla 309
Oy 955 GTGGTTGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1014
    |||||.....|
Db 310 ValValAlaArgGlyGluPhePheProArgHsProGlyLeuLeuAlaSerGluAerPhe 329
Oy 1015 AAAGCAATTCCTTCATCATGAGAGTCAATTAACAAGAGTGGCTTCTGCTGCTATG 1074
    |||||.....|
Db 330 HsProValProSerIleIleGlyValAlaAsnAerGluPheGlyTrpThrIlePro--- 348
Oy 1075 AAGAGAGCTCTGAGATCTCTGAGTGGCTTCCACAAGTCCCTTGGCCCTCATCTGATACA 1134
    |||||.....|
Db 349 -----ValValMetGlySerAlaGlnTrpIleGlyGluIleThrArgGlu 363
Oy 1135 AACATC-----CTGCACATCCCGCTCAGATATTG 1164
    |||||.....|
Db 364 AsnLeuGlnAlaValLeuLeuAerhTrhAlaGlnIleMetLeuProProGlyCySer 383
Oy 1165 CACCTTGTGGCTAATGAATTAATCTTCATGAACAAGACTCCCTGATCTGAATTCGAGACT 1224
    |||||.....|
Db 384 AspLeuLeuMetGluGluTyMetGlyAerhTrhGluAerAlaGlnIleThrLeuGlnIleGln 403
Oy 1225 CTTCTGAGACTGCTGGAGAGTGTGTTCTTGTGGTCCCTGCTGATCATCAGCTGATAT 1284
    |||||.....|
Db 404 PhehTrhGluMetMetGluAerPheMetPheValIleProAlaLeuGlnValAlaTyPhe 423
Oy 1285 CACAGAGATGCTGGAGACTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1344
    |||||.....|
Db 424 GlnArg---SerhAlaSerValTyPheTyTrpGluPheGlnIleAlaSerLeu 442
Oy 1345 GAAGAACAAGACCGGCTTTTGTCAAGCCGACACAGCTGATGAAGTCCGCTTGTGTT 1404
    |||||.....|
Db 443 TyAerValArgProTrhIleValTyAlaAerhAlaAerGluIleProhValPhe 462
Oy 1405 GGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1458
    |||||.....|
Db 463 GlyTyPhePheTrpAerMetTyLeuAerPhe-----ThrGluGly 476
Oy 1459 GAGAAAGTACTGAGCGCGAGATGAATATCTGGGCTTCTTCTTCTTCTTCTTCTTCTTCTT 1518
    |||||.....|
Db 477 GlnTyLeuLeuSerArgTrhMetCysTyTrpAlaAerhAlaAerhAlaArgHsGlyAsn 496
Oy 1519 CTTAATGGAGACGACTGTCTGTGCGCAGCTTAAATCTGATCTGAGAGTACTTCCAG 1578
    |||||.....|
Db 497 ProAerSerGlnGlyLeuProTyTrpProValMetAerhAlaAerGlnIleTyLeuGln 516
Oy 1579 CTGAGACTTGAACATGAGCTCGAGACAGACTCAAGAAACCGCGGTGATTTTGAGAC 1638
    |||||.....|
Db 517 LeuAerTrhGlnProAlaValGlyAlaLeuTySerArgTrhGlnIleThrPhe 536
Oy 1639 AGCAACATCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1689
    |||||.....|
Db 537 TyhTrhLeuSerGlnTyLeuIleGlnIleAlaSerGlnAerPheTyhTrhLeu 556

```

RESULT 14

```

Oy 06GMS4 PRELIMINARY: PRT: 568 AA.
Db 06GMS4:
Oy 05-JUL-2004 (TrEMBLrel. 27, Created)
Db 05-JUL-2004 (TrEMBLrel. 27, last sequence update)
Db 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
Db LOC43703 protein (Fragment).
GN Name=LOC43703;

```

```

Oy 05 Xenopus laevis (African clawed frog).
Oy Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Oy Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Oy Xenopodinae; Xenopus.
Oy NCBI_TaxID=8355;
Oy (1)
Oy SEQUENCE FROM N.A.
Oy TISSUE=Kidney;
Oy MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Oy Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Oy Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
Oy Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Oy Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Oy Diatchenko L., Maruina K., Farmer A.A., Rubin J., Hong L.,
Oy Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
Oy Brownstein M.J., Uedlin T.B., Tothylaki S., Carrinci P., Prange C.,
Oy Kane S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Oy Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Oy Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Oy Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Oy Fahy J., Hellon E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
Oy Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Oy Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Oy Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
Oy Krzywinski M.I., Skalski U., Smalins D.E., Schnerch A., Schein J.E.,
Oy Jones S.J., Maira M.A.;
Oy "Generation and initial analysis of more than 15,000 full-length human
Oy and mouse cDNA sequences."
Oy Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
Oy (2)
Oy SEQUENCE FROM N.A.
Oy TISSUE=Kidney;
Oy MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
Oy Klein S.L., Struhsberg R.L., Wagner L., Pontius J., Clifton S.W.,
Oy Richardson P.;
Oy "Genetic and genomic tools for Xenopus research: The NIH Xenopus
Oy Initiative."
Oy Dev. Dyn. 225:384-391 (2002).
Oy (3)
Oy SEQUENCE FROM N.A.
Oy TISSUE=Kidney;
Oy Klein S., Gerhard D.S.;
Oy Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
Oy -1- SIMILARITY: Belongs to the type-B carboxylesterase/1lipase family.
Oy EMBL, BC074230; AA074230.1; -.
Oy GO; GO:0016787; F:Hydrolase activity; IEA.
Oy InterPro; IPR002018; Carboxylesterase.
Oy InterPro; IPR00379; Ser_esterase.
Oy Pfam; PF00135; Coesterase; 1.
Oy PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
Oy PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
Oy Hydrolase.
Oy FT NON TER 1 1
Oy SQ SEQUENCE 568 AA; 63032 MW; 9336D0B8B53931A1 CRC64;

```

Alignment Scores:

```

Pred. No.: 2.64e-85 Length: 568
Score: 1263.50 Matches: 258
Percent Similarity: 64.49% Conservative: 87
Best Local Similarity: 48.22% Mismatches: 169
Query Match: 39.66% Indels: 21
DB: 2 Gaps: 8

```

US-10-023-515-3 (1-1746) x 06GMS4 (1-568)

```

Oy 82 GAAAAAAGTGGGCTTGTCTGTAAGGCGCACAGAGAACACGAGCTGAGATTCAG 141
    |||||.....|
Db 28 GlyIleTyArgIleTrhGluAerAlaArgProLeuLeuTrhTrhAerTyGlyGlnLeuLeu 47
Oy 142 GCGAAGATGTCATCTGCTGCTGAGAGCCCTGTGCTGTAACGTTTCTTCTGAGTCCCT 201
    |||||.....|
Db 48 GlyTyhTrhValGlyAlaTyLeuGluTrhAerArgLeuIleHsValPheMetGlyValPro 67

```

```
OY 202 TTGCTGCTCCCGCTGGAGATCCCTGCATTTACGAACCCGCGAGCTGCATCGCCCTG 261
Db PheAlaIyPProIleGlyProLeuAArgPheGluAAsPProGInProGluPPro 87
OY 262 GATTAATTCGGAAGACCCATCTCTACCTTAATTGCTCTCCAGAACATCAAGATGGCTG 321
Db SerSerIleArgGluAlaThrGluAAsnPProPromeCyLeuGluAAsPylAsySglYmet 107
OY 322 -----CTCTGATCAACACATGCTCAAGGTGCANTTACCGGAANTCGAGTGCAGAA 375
Db GluGInLeuAlaAsPPhenPheIyAlaIyAsPPhenAAsPPro-----ValSerGlu 125
OY 376 GACTGCTCTTACCTGAACATCTATGCGCTGCGCCACGCGCATACAGCTCCAGCTCCCC 435
Db AAsPylAsyLeuAAsnPAlaPheThrProAlaAAsPArgGlyGluAAsnPProIleuPro 145
OY 436 GCTTTGGTGTGGTCCCGAGAGGTGCTTCAAGACTGCTCAAGCTCCATCTTTGAATGG 495
Db ValMetValPheIleHISglYglYleuThrMetGlyAlaGlyMetPheGluGly 165
OY 496 TCCGCGCTGGCTGCTATGAGGACGTGCTGTTGTGTGCTGCTGACATCCGCTAGGAATA 555
Db SerAlaLeuSerAlaArgGluAAsnPAlaValValSerIleGlnTyArgLeuGlyIle 185
OY 556 TTTGGTTCTTCACACATGGGATCAGCATGCTCCGCGGAACTGGAGCTTCAAGACAG 615
Db MetGlyPhePheSerThrGlyAAsPylAsyAlaArgGlyAAsnTyGlyPheIleuAAsPIn 205
OY 616 GTGCTGCTCTGCTCTGGGTCCAGAAACATCGACTTCTTGGTGGGAGCCCAAGCTCT 675
Db ValAlaAlaLeuAlaArgTyAlaArgAAsnPleTyAAsPPhenGlyIAsnPProInSer 225
OY 676 GTGACATCTTTGGGAGTCCGCGGAGCCATAGCTTTTCTTCACTTATCTGTCTCC 735
Db ValThrIlePheGlyGluSerAlaGlyIleuSerValSerAlaGlnValIleuSerPro 245
OY 736 ATGGCCAAAGGCTTATTCACAAACCATCATGAGAGTGGGTGGCCATCATCCCTTAC 795
Db LeuSerIySglYleuPheHISArgAlaIleSerGluSerGlyValAlaIleuProGly 265
OY 796 CTGAGAGCCCATGATTATGAGAAAGTGAAGCTGCAGAGTGGTTGCACATTTCTGTG 855
Db LeuMetAla-----SerIyThrGluIySglIleuProIleuHISAlaValAla 282
OY 856 AACATGCGTCAAGACTCT--GAGGCTGCTGAGAGTGCCTGAGGACAAACCCCTCAAG 912
Db AAsnIleSerSerCySerValSerSerIleuAlaAsPylAsyLeuTyThrGluAAsP 302
OY 913 GAGTGTGACCTTCAGCCACAGAAACAAAG-----TCTTTCATCTCGAGTGGTATGGT 966
Db GluIleValAlaIleSerAlaAlaMetIySphValAlaPheProAlaValIAsPgly 322
OY 967 GCTTCTTCTTAATGAGCTCTAGACTTATTTGTCTGAAAGCATTTAAAGCAATTCCT 1026
Db ValPheIleuProIySProAlaGluGluIleuAAsnIySglIleuAAsnPProValPro 342
OY 1027 TCCATGCTGAGTCAATTAACACAGAGTGGCTTCCGCTGCGCTTGAAGAGAGCTTCCT 1086
Db PheIleuIleGlyValAAsnPAlaHISglYIleuPheGlyIleuProIleuAlaIleuAAsnP-- 361
OY 1087 GAGATCTCAGTGGCTCCACAAAGTCCCTTGCTCCATCTGATACAAACATCTCGAC 1146
Db -----IleSerGlyTyArgGluGlyMetGlyIySAsPylIleGlnSerIleVal 379
OY 1147 ATCCCGCTCAGTATTTGCAC-----CTTGGGCTTAATGAATAC 1185
Db AlaLeuPro--PheValHISerPheThrSerValIProPheIleMetGluGlyTy 398
OY 1186 TTCCATGACAAAGCATCCCTGCTGAATTCGGAATCTTCTGAGCTTCTGGAGAT 1245
Db PheGlyAAsPThrAAsnPProIySglIleuAArgAAsnPheIleuAAsPLeuValGlyAAsP 418
OY 1246 GTGTTCTTTGTGTCCCTGCATGATCAAGCTCGATATCAAGAGATGCTGTGCACT 1305
```

```
Db 419 IleIlePheValIleProAlaIleuAArgThrAlaIySerTyHISArgAAsPSeGlyHISPro 438
OY 1306 GTTACTTCTTAATGAGTTTGGACCGGCTCGAGTCTTTGAAGACACAGACCGGCTTT 1365
Db ValTyPheTyArgIlePheGlnHISArgPProSerMetTyAAsPAsPSeTyAAsPAsPhe 458
OY 1366 GTCAAGCCGACACCGCTGAGTGAAGTCCGCTTGTGTTGGTGGTGGCTTCTGAAGGG 1425
Db ValIySAlaAsPAlaHISglYAsPglIleuTyPheValAlaGlyIyProPheIleuIySer 478
OY 1426 GACATTTGATGTTGCAAGGACGACGAGGAGAGAGAAATTACATGAGCCGGAATGATG 1485
Db GlyIleuPheIySAsPAsnGlyThrGluGluIySleIleuSerIySThrIleMet 498
OY 1486 AATATCTGGGCTACCTTTGCTCGAACCGGAAATCTTAATGGGAAAGCATGCTCTGTGG 1545
Db LysTyTrpAlaAAsnPheAlaArgAAsnGlyAAsPProAAsnGlyLeuAlaGlyIleu 518
OY 1546 CCAGCTTATATCTGACTGACATGACTTCCAGCTGACATCTGAAACATGAGCTCGACAG 1605
Db ProIyTyAAsPglIuAAsPglIuAAsPtyIleuGluIleIySleuThrGlnIySleuSerGln 538
OY 1606 AGACTCAAGAAACCGGCTGAGTTTGGACGACACCATCCCC 1650
Db ArgIleuIySglYArgPheIyPheThrPThrValThrIleuPro 553

RESULT 15
ID 070177 PRELIMINARY; PRT; 561 AA.
AC 070177;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DB Carboxylesterase precursor (BC 3.1.1.1).
GN Name=carboxylesterase;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Small intestine;
RA Sone T., Kunimoto T., Isobe M.;
RU Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.
DR EMBL; AB010632; BAA25691.1; -.
DR HSBP; P37967; I0E3.
DR GO; GO:0004091; F:carboxylesterase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR002018; CarbesteraeB.
DR InterPro; IPR00379; Ser_estrs.
DR Pfam; PF00135; Coesterase; 1.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolase; Signal.
FT SIGNAL 1 26 Potential.
FT CHAIN 27 561 carboxylesterase.
SQ SEQUENCE 561 AA; 62239 MW; 73A468C3F96939B CRC64;

Alignment Scores:
Pred. No.: 3,12e-85 Length: 561
Score: 1262.50 Matches: 275
Percent Similarity: 62.41% Conservative: 77
Best Local Similarity: 48.76% Mismatches: 179
Query Match: 39.83% Indels: 33
DB: Gaps: 13

US-10-023-515-3 (1-1746) x 070177 (1-561)
OY 1 ATCCCAAGGGA--CTTACTCATCTGCTTCAAAATGAGCTTTTCTGATTCGAC 57
Db 1 MetProArgAAsnPInIleuHISerTrPLeuAAsnPAlaValIleuPheGlyIleuIleu--- 19
```

```

Oy 58 CCCCTGTTGGACACAGACAGTGGGAAAACTGGGCTTCTGTGAAGGCCACAGAG 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20 ---LeuLeuIleHisValGln---GlyGlnAspSerProGluSer---SerProIleArg 36
Oy 118 AACACGAGCTGGAGTGGATTCAGGGACGACAGTCACTGTCTGGAGCCCTTGCT 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 ThrTrhHisThrGlnGlnValGlnGlnIlyLeuAspHisIleValIyAspThrIyValIagly 56
Oy 178 GTGAACGTTCTCTGGAGAGTCCCTTTTGCTGTCCCCCGCTGGAGATCCCTGGATTAG 237
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 57 ValHisThrPheLeuGlyIleProPheAlaIyAspProValGlyProLeuAspPheAla 76
Oy 238 AACCCGAGCTGCATCGCCCTGGATTACTGCAGAGACCACTCCCTCAATTTTGG 297
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 ProProGluProProGluProIyPheSerGlyValIyAspAlaHisSerGlnProAlaMet 96
Oy 298 TGCCTCCAGAACTCAGAGTGGCTG-----CTTTAGATCAACATGCTCAAG 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 CysLeuGlnIleAsnLeuAspIleLeuAspGlnValGlyLeuLeuAspMetIyMetIle--- 115
Oy 346 GTGCATTACCCGAATTGGAGATGCAGAGATCCCTCTACCTCAATCATCTAGCTGCT 405
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 116 ----LeuSerSerIleSerMetSerGluAspCysLeuIyIyLeuAsnValIyIyAlaPro 133
Oy 406 GCCCAGCCGATACAGGCTCCAGGCTCCCGCTTGATGTTGGTCCAGAGAGCTCTTC 465
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 134 AlaHisAlaAspGlnGlyIySerAsnLeuProValIyMetValIyIyPheIleGlyIyAlaLeu 153
Oy 466 AAGATGCTCAGCTCCATCTTTGATGGGTCCGCTGGCTGCTCATAGAGACGTGCTG 525
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 ValValGlyMetAlaSerMetIyIyAspGlySerLeuLeuThrValIyAsnGlnAspLeuVal 173
Oy 526 GTTGTGTCGTCACTGATCCGGCTAGAGATATTGGTTCTTCAACAATGGATCAAGAT 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 ValValIyIyIyGlnIyIyArgLeuGlyValLeuGlyPhePheSerThrGlyAspIyIyHis 193
Oy 586 GCTCCGGGGAAGCTGGGCTCAAGGACGAGGCTGCTGCTGGGCTCCAGAAAGAC 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 194 AlaArgGlyAsnTrpGlyIyIyLeuAspGlnValAlaAlaLeuAspIyIyIyValGlnIleAsn 213
Oy 646 ATCGAGTCTTCTGGTGGGACCCCAAGCTCTGTGACCATCTTTGGAGAGTCCGCGGAGCC 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 214 IleAlaHisPheGlyGlyAsnProAsnArgValThrIlePheGlyGluSerIleGlyIy 233
Oy 706 ATAACTTTTCTAGTCTTAATCTGTCTCCATGCGCAAGGCTTATTCACAAAGCATC 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234 ThrSerValSerSerHisIleValIleSerProMetSerGlnGlyLeuPheHisIleGlyAlaIle 253
Oy 766 ATGAGAGTGGGGTGGCCATCATCCCTTACCTGAGAGGCCCATGATTAGAGAAAGTAG 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 254 MetGluSerGlyValAlaLeuLeuProAspLeuIleSer-----GluThrSerGlu 270
Oy 826 GACCTGCAG---GTGGTGCACATTTCTGTGTAAACAATGCGTCAGACTGTAGGCCCTG 882
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 271 ThrValSerThrThrValAlaIyAspLeuSerGlyCysGlnAlaMetAspSerGlnAlaLeu 290
Oy 883 CTGAGAGTGCCTGAGGACAAACCTCCAAAGAGCTGTGACCCCTGACCCAGAAACAAAG 942
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 291 ValArgCysLeuAspGlnAlaIySerGlyAlaGlnIleLeuValIleAsnIyValPheIyAs 310
Oy 943 TCTTTCACTCGAGTGGTGAATGTGCTTTCTTTCTTAATGAGCCTTAGATCTATGTCT 1002
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 311 MetIleProAlaValValAspIyGlnPheLeuProAlaArgHisProIyAspIyLeuLeuAla 330
Oy 1003 CAGAAAGCATTTAAAGCAATCTCTTCAATCATGAGATCAATACCAACGAGTGGGCTTC 1062
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 331 SerGluAspPheHisProValProSerIleIleGlyValAsnThrAspGluIyIyCysCys 350
Oy 1063 CTGCTGCTTANG-----AAGAGAGCTCTGAGATCTCACT 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 351 ThrIleProMetValMetGlyThrAlaGlnIleIleIyGln-----LeuSer 366

```

```

Oy 1099 GGCCTCAACAAATCCCTTGCCCTCCATCTGAATACAAACATCTGCATCCCGCTCAG 1158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 ArgGluAsnLeuGlnAlaValIleLeuAspThrAlaAlaGlnMetMetLeuProProGlu 386
Oy 1159 TATTTCACCTTGGCTGAATGAATATCTTCATGACCAAGACATCCCTGACTGAAATCCGA 1218
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 CysGlyAspLeuLeuMetGluGlnIyIyIyMetGlyAsnThrAspAspSerGlnThrLeuGln 406
Oy 1219 GACAGCTTTCTGAGCTTCTGAGAGATGTGTTCTTTGTGGTCCCTGCATGATCAACACT 1278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 IleGlnIyIyIyIyIyIyMetMetGlyAspPheLeuPheValIleProAlaLeuGlnValAla 426
Oy 1279 CGATATCAACAGATGCTGTGACCTGTCTATCTTATGAGTTTGGCAGCCGCTCAG 1338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 HisPheGlnArg---SerHisAlaProValIyIyPheIyIyGlnPheGlnHisAlaProSer 445
Oy 1339 TGCCTTGAACACAGAACCCGCTTTTGTCAAGCCGACCAACGCTGATGAATGCCCTTT 1398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 446 TyrPheIyAsnValIyArgProProHisIleValIyAlaAspHisAlaAspGlnValIyProHe 465
Oy 1399 GTGTTGCGTGGTGCCTTCTGAGAGGGGACATGTTGTTATGTTGGAAGAGCCAGAGAG 1458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 ValPhe---GlySerPhePheSerGlyMetIyIyLeuAspPhe-----ThrIyGln 481
Oy 1459 GAGAGTTACTGAGCCGAGAGATGATGAATATCTGGGCTTACTGTTGCTGAACCGGGAAT 1518
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 GlnArgLeuLeuSerArgIyAspMetMetIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 501
Oy 1519 CCTAATGGAAACGACCTGTCTGTGGCCATTAATCTGACCTGAGACAGTACTCCAG 1578
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 ProAsnSerGlnGlyIyLeuProIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 521
Oy 1579 CTGACCTTGAACATGAGCTGTGAGACAGAGCTCAAGAAACCGGGGTGAGATTTTGGACC 1638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 522 LeuAspThrHisProAlaValAspArgAlaLeuIyValIyAlaArgIyLeuGlnPheThrIy 541
Oy 1639 AGCACCATCCCC 1650
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 LysThrIleuPro 545

```

Search completed: June 15, 2005, 08:53:13
Job time : 183.861 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: June 15, 2005, 07:48:14 ; Search time 26.3868 Seconds
(without alignments)
9878.981 Million cell updates/sec

Title: US-10-023-515-3
Perfect score: 3170
Sequence: 1 atgcacagagactactctc.....cttctcttgcctcttga 1746

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p model -DEV=x1h
-O=/cgn2_1/USPTO_spool/6664091/runal_14062005_133500_9360/app_query.fasta_1.4238
-DB=Issued_Patents_AA -QFMT=fastcan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -NOR=ext -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=psco -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=6664091 @cgn2_1_154 @runat_14062005_133500_9360 -NCPU=6 -ICPU=3 -NO_MMAP
-LARGEQUERY -NSG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRAADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pcp:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pcp:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pcp:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pcp:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pcp:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3079	97.1	581	4	US-10-023-515-2
2	1230	38.8	559	4	US-09-595-682B-28
3	1230	38.8	559	4	US-09-949-016-6426
4	1230	38.8	577	4	US-09-949-016-9670
5	1150.5	35.3	565	4	US-09-595-682B-21
6	1138.5	35.9	566	3	US-09-264-737-2
7	1130.5	35.7	543	4	US-09-595-682B-26
8	1091.5	34.4	584	2	US-08-845-295A-2
9	1091.5	34.4	584	3	US-09-140-933-2
10	1091.5	34.4	584	3	US-09-146-661-2
11	1091.5	34.4	584	3	US-09-150-515-2
12	1034.5	32.6	539	3	US-09-264-737-1

13	942.5	29.7	454	3	US-08-446-100-26	Sequence 26, Appl
14	942.5	29.7	454	3	US-08-446-100-28	Sequence 28, Appl
15	942.5	29.7	454	3	US-08-446-100-30	Sequence 30, Appl
16	942.5	29.7	454	3	US-08-446-100-31	Sequence 31, Appl
17	940.5	29.7	454	3	US-08-446-100-27	Sequence 27, Appl
18	940.5	29.7	454	3	US-08-446-100-29	Sequence 29, Appl
19	871	27.5	574	4	US-10-023-515-4	Sequence 4, Appl1
20	793	25.0	933	4	US-09-949-016-8386	Sequence 8386, Ap
21	786	24.8	823	4	US-09-949-016-6888	Sequence 6888, Ap
22	782	24.7	953	4	US-09-949-016-8387	Sequence 8387, Ap
23	777.5	24.5	843	4	US-09-491-356C-20	Sequence 20, Appl
24	777	24.5	575	1	US-08-348-920-1	Sequence 1, Appl1
25	775	24.4	823	4	US-09-491-356C-23	Sequence 23, Appl1
26	774	24.4	575	1	US-08-348-920-2	Sequence 2, Appl1
27	772	24.4	848	4	US-09-491-356C-22	Sequence 22, Appl1
28	769	24.3	836	4	US-09-491-356C-21	Sequence 21, Appl
29	765.5	24.1	614	3	US-08-446-100-25	Sequence 25, Appl
30	764.5	24.1	614	1	US-07-732-962A-2	Sequence 2, Appl1
31	764.5	24.1	614	2	US-08-370-156-2	Sequence 2, Appl1
32	764.5	24.1	614	3	US-08-446-100-19	Sequence 19, Appl
33	764.5	24.1	614	3	US-08-446-100-21	Sequence 21, Appl
34	764.5	24.1	614	3	US-08-814-095-2	Sequence 2, Appl1
35	764.5	24.1	614	5	PCT-US92-06106-2	Sequence 2, Appl1
36	764.5	24.1	645	4	US-09-949-016-7063	Sequence 7063, Ap
37	764.5	24.1	645	4	US-09-949-016-7064	Sequence 7064, Ap
38	762.5	24.1	614	3	US-08-446-100-20	Sequence 20, Appl
39	761.5	24.0	614	3	US-08-446-100-23	Sequence 23, Appl
40	760.5	24.0	614	3	US-08-446-100-22	Sequence 22, Appl
41	759.5	24.0	600	2	US-08-370-156-4	Sequence 4, Appl1
42	759.5	24.0	600	3	US-08-814-095-4	Sequence 4, Appl1
43	759.5	24.0	600	3	US-08-975-084-1	Sequence 1, Appl1
44	755.5	23.8	617	2	US-08-370-156-6	Sequence 6, Appl1
45	755.5	23.8	617	3	US-08-814-095-6	Sequence 6, Appl1

ALIGNMENTS

```
RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Patent No. 6664091
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: 53010. A NOVEL HUMAN CARBOXYL ESTERASE
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023,515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279,508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-515-2

Alignment Scores:
Pred. No.: 2,86e-294
Score: 3079.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 97.13%
DB: 4
Gaps: 0
US-10-023-515-3 (1-1746) x US-10-023-515-2 (1-581)
1 ATGCACAGGAGACTTCTCATCTGCTTCACATGAGTGGCTTTTCCTGATTCACAGCC 60
|||||
```

Db 1 MecProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuIleuGlnPro 20
QY 61 CTGTTGGGACACAGACAGTGGGAAAACTGGGCTTCTGCTGAAGGGCCACAGAGAAC 120
Db 21 LeuLeuGlnYHhSARGlnTrpGlyThrGlyProSerAlaGlnGlyProGlnTrpGln 40
QY 121 ACCAGGCTGGAGTATTCAGGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGGCTGTG 180
Db 41 ThrArgLeuGlnGlyTrpIleGlnGlySerGlnValThrValLeuGlySerProAlaProVal 60
QY 181 AACGTTTCTCTGGAGTCCCTTTGTGCTGCCCGCTGGAGTCCCTGGATTACGAAC 240
Db 61 AsnValPheLeuGlyValProPheAlaAlaProLeuGlySerLeuAlaArgPheThrAsn 80
QY 241 CCGGACCTGGATGGCCCTGGGATACTTGGAGAAAGCACTCTTACCTTAATTTTGGC 300
Db 81 ProGlnProAlaSerProTrpAsnLeuArgGlnValaThrSerTrpProAsnLeuCys 100
QY 301 CTCGAGAACTCAGAGTGGCTGCTTGAATCAACACATGCTCAAGGTGATTAACCGAAA 360
Db 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHISMetLeuLeuValHISerTrpGlyS 120
QY 361 TTCGAGTGTCAAGAACTGCTCTAAGCATCTATGCGCTGCCACGCGCATACA 420
Db 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHISAlaAspThr 140
QY 421 GGCTCAAGCTCCCGCTGTGGTGTGTTCCAGAGAGGTGCTTCAAGATGGCTCAGCC 480
Db 141 GlySerIleLeuProValLeuValTrpPheProGlyGlyAlaPheLeuTrpHISerAla 160
QY 481 TCCATCTTGTAGTGGTCCGCCCTGGCTCTATAGAGACGTGCTGTGGTGTGCTCAG 540
Db 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValGln 180
QY 541 TACCGGCTAGAAATTTGGTGTCTTCAACACATGGGATCAAGCATGCTCCGGGAACTGG 600
Db 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHISAlaProGlyAsnTrp 200
QY 601 GCCTTCAAGACAGAGTGGCTGTCTGCTGGGTCCAGAAACATCGATTTCTTCGGT 660
Db 201 AlaPheIleAspGlnValAlaAlaLeuSerTrpValGlnIleAsnIleGlnPhePheGly 220
QY 661 GGGGACCCAGCTCTGTGACCATCTTTGGGAGTCCGGGAGCCATTAAGTGTCTTAGT 720
Db 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaIleSerValSerSer 240
QY 721 CTTATACGTGTCCCAAGCCCAAGGCTTATCCCAAGCAAGCATCAGAGAGTGGGTG 780
Db 241 LeuIleLeuSerProMetAlaIleGlyLeuPheHISValAlaIleMetGlnSerGlyVal 260
QY 781 GCCATCATCCCTTACCTGAGGCGCATGATTATGAGAAGTGAAGACCTGACGGTGGT 840
Db 261 AlaIleIleProGlyTrpLeuGlnAlaHISAspTrpGlyIleuSerGlnAspLeuGlnValVal 280
QY 841 GCACATTTCTGTGTAAACAATGCGTCAAGCTTGAAGCCCTGTGAGGTGCTTGAAGACA 900
Db 281 AlaHISpHeCysGlyAsnAlaAlaSerAspSerGlnAlaIleuLeuArgCysLeuArgThr 300
QY 901 AAACCCCTCAAGAGCTGCTGACCCCTCAGCCCAAGAAAACAAGTCTTTCACCTGAGTGGT 960
Db 301 LysProSerIleGlyGlnLeuLeuThrLeuSerGlnIleTrpLysSerPheThrArgValVal 320
QY 961 GATGATCTTTCTTCTCTAATGAGCTCTAATGATCTATGTTCTCAGAAAGCATTTAAAGCA 1020
Db 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnIleValPheLeuVala 340
QY 1021 ATTTCCTTCATCATCGAGTCAATTAACCAAGAGTGTGGCTTCTGCTGCTATGAAGAG 1080
Db 341 IleProSerIleIleGlyValAsnAlaHISGlnCysGlyPheLeuLeuProMetLeuGln 360
QY 1081 GCTCCTGAGATCCTCAGTGGGTCCACAAGTCCCTTGGCCCTCAATGATCAAAACATC 1140
Db 361 AlaProIleIleLeuSerGlySerAsnLysSerLeuAlaLeuHISLeuIleGlnAsnIle 380

QY 1141 CTGACATCCGCGCTCAGATATTTGACCTTGGCTAATGAATCTTCATGACAGAGAC 1200
Db 381 LeuHISIleProProGlnTrpLeuHISLeuValAlaAsnGlnTrpPheHISAspLysHIS 400
QY 1201 TCCCTGACTGAATCCGAGACAGTCTTCTGGAATTTGCTTGGAGATGTGTTCTTGTGCTC 1260
Db 401 SerLeuTrpGlnIleArgAspSerLeuLeuAspLeuLeuGlnLysAspValPhePheVala 420
QY 1261 CCTGACATGATCAAGCTTCGATATCAACAGAGATGTGTGACCTGTCTTACTTCTATGAG 1320
Db 421 ProAlaLeuIleTrpAlaArgTrpHISArgAspAlaGlyAlaProValTrpPheTrpGln 440
QY 1321 TTTGGCACCGGCGCTCAGTGTCTTGAAGACACAGAGCCGGCTTTTGTCAAAAGCCGACAC 1380
Db 441 PheArgHISArgProGlnCysPheGlnAspTrpLysProAlaPheValLysAlaAspHIS 460
QY 1381 GCTGATGAAGTCCGCTTGTGTGCTGTGCTGCTTCTGAAAGGGGACATTTGATGTTTC 1440
Db 461 AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
QY 1441 GAAGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGAAATATCGGCTTACC 1500
Db 481 GlnGlyAlaTrpGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpAlaThr 500
QY 1501 TTTGCTGGAACCGGGAATCCTAATGGGAAAGCACTGTCTCTGTGGCCGCTTATTAATCTG 1560
Db 501 PheAlaArgTrpGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
QY 1561 ACTGAGACGATCCTCCAGCTGGACTTGAACATGAGCTTCGACAGAGACTCAAGAACCG 1620
Db 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlnArgLeuLysGlnPro 540
QY 1621 CGGGTGATTTTGGACACAGACATCCCTGATTCCTGTGTGCTCGCATATGCTCCAC 1680
Db 541 ArgValAspPheTrpTrpSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHIS 560
QY 1681 AGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCAGACCTTCTTCTTCTTGTGCT 1740
Db 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
QY 1741 CCT 1743
Db 581 Pro 581

RESULT 2
US-09-595-682B-28
/ Sequence 28, Application US/09595682B
/ Patent No. 6800483
/ GENERAL INFORMATION:
/ APPLICANT: Danke, Mary K.
/ APPLICANT: Potter, Phillip M.
/ APPLICANT: Houghton, Peter J.
/ TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
/ FILE REFERENCE: SJ-0005
/ CURRENT APPLICATION NUMBER: US/09/595,682B
/ PRIOR APPLICATION NUMBER: 60/075,258
/ PRIOR FILING DATE: 1998-02-19
/ PRIOR APPLICATION NUMBER: PCT/US99/03171
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 28
/ LENGTH: 359
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-595-682B-28

Alignment Scores: 4.44e-112 length: 559
Pred. No.: 1230.00 Matches: 248
Score:

Percent Similarity: 63.84% Conservative: 91
 Best Local Similarity: 46.70% Mismatches: 168
 Query Match: 38.80% Indels: 24
 DB: 4 Gaps: 8

US-10-023-515-3 (1-1746) x US-09-595-682B-28 (1-559)

```

QY 91 GGGCTCTGGAGAGGGGCAAGAGGAAACACAGGCTGGAGTTCAGGGCAAGAA 150
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 GTCACTGTCTGGAGAGGCTGTGCTGTGAACGTGTCTTGTGAGTCCCTTGTCT 210
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 46 ValHisValIyVgIyAlaAsnIaIyValIghIthrPheIuGlyIleProheIaIy 65
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 CCCCCGTGGAGTCCCTGGAGTTTACGAACCCGAGCTGGACCCCTGGAGTAACTTG 270
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 66 ProProIeuGlyIProIeuIaIgrPheIaIProProGluIProGluSerItrPserGlyVal 85
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 CGAGAGGCACTCTACCTTAATTGTGCTCCAGAAC-----TCAGAG 315
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 86 ArgSerGlyIThrThrIaIProIaIaMetCyIeuGlnAerIuThrAlaValGluSerGlu 105
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 TGGCTGTCTTAGATCAACACATGCTCAAGGTGATTAACCCGAATTGGAAGTGCAGA 375
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 106 PheIeu-----SerGlnPheAsnMetIthrPheProSerIaPserMetSerGlu 121
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 GACTGCTCTTACATGAACATGTATGCGCTGCGCCAGCCGAGTACAGCTCCAGCTCC 435
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 122 AaPcyIeuIyIleuSerIeIyIthrProIaIaIserIaIeGluIySerAenIeuPro 141
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 GTCTTGTGTGTGTCCCAAGAGTGTCTTCAAGAGTGTGCTGACCTGATCTTGAATGG 495
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 142 ValMetValItrIleIaIeGlyIAlaIeuValIaPheGlyMetAlaSerIeuIyIaPrg 161
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 TCCGCGCTGGCTGCTATGAGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 162 SerMetIeuAlaIaIaIeuGlnAenValIaValIleIleGlnIyIaPrgIeuGlyAl 181
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 TTTGGTTTCTTCAACACATGGAGTACAGATGTCTCCGGGAGCTGGACCTTCAAGACAG 615
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 182 LeuGlyIaPhePheSerIthrGlyAaPryIaIaIaIaIaIaIaIaIaIaIaIaIaIa 201
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 616 GTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 675
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 202 ValAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 221
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 676 GTGACATCTTTGGAGAGTCCGCGGAGCATTAAGTGTCTTGTGTGTGTGTGTGTGT 735
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 222 ValThrIlePheGlyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 241
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 736 ATGGCAAGAGGCTTATTCACAAAGCCATGATGAAGTGGGGTGGCCATCCCTTAC 795
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 IleSerGlnIyIleuPheIaIeGlyIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 261
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 796 CTGAGAGCCCATGATTAAGAGAGTGAAGCTG-----CAGTGTGTCAATTTTC 849
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 262 LeuIleIaIa-----SerSerIaAaPryValIleSerIthrValaIaIaIaIaIe 277
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 850 TGTGTGAACATAGGCTCAGACTCTGAGGCGCTGAGAGTGTGTGTGTGTGTGTGTGT 909
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 278 SerAlaCyAaPryAaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 297
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 910 AAGAGTGTGTGACCTTCAAGCCAGCAAAACAAAGTCTTTCATCGAGTGTGTGTGTGT 969
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 GluGluIleIeuAlaIleAenIyAaPryProPheIyMetIleProGlyValaIaPrgIyVal 317
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 970 TTTCTTCTTAATGAGCTTAAATCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1029
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 318 PheIeuProIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 337
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1030 ATCATGAGTGAATTAACACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1077
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

DB 338 IleValIyValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 357
QY 1078 GAGGCTCTGAGATCTCAAGTGGCTTCCAAAGTCCCTTGGCTTCATCTGATCAAAAC 1137
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 358 AaPThrGlnIyVgIuMetAaPryGlyAlaSerGlnAlaIaIaIaIaIaIaIaIaIaIaIa 377
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1138 ATCTGCAATCTCCGCTCAAGTATTTGCACTTGTGGTGAATGAATCTTCAATGCAAG 1197
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 378 LeuIeuMetIeuProProIthrPheGlyAaPryIeuAaIaIaIaIaIaIaIaIaIaIaIa 397
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1198 CACTCCCTGATGAAATCCGAGACAGTCTTGTGAGCTTGTGTGAGATGTGTCTTTGG 1257
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 GlyAaPryGlnIthrIleuGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 417
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1258 GTCCCTGCAATGATACAGCTGATGATACAGAGATGTGGTGTGCACTGTCTTCTTAT 1317
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 418 IleProIaIaIeGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 436
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1318 GAGTTTGGGCAACCGGCTCAAGTGTGTGTGAAGACAGAAAGCCGCTTGTGTCAAGCCGAC 1377
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 437 GluPheGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 456
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1378 CAGCTGATGAATCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1437
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 457 HisGlyAaPryGluIeuProPheValPhe--ArgSerPhePheGlyGlyAaIyIaIy 475
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1438 TTGCAAGAGACCAAGGAGAGAGAAATTAAGTACAGCCGGAAGATGAATTAAGTGGCT 1497
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 476 Phe-----ThrGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 492
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1498 ACCTTGTCTGAAACCGGAAATCTTAATGGAAGACGACTGTCTGTGTGCACTTATTAAT 1557
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 493 AaPheIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 512
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1558 CTGACTGAGCACTTCACTCAAGTGTGAATCAATGAGCTTGTGCAAGACTCAAGAA 1617
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 513 GlnGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 532
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1618 CCGCGGTGTGATTTTGTGACCAAGCAACATCCCC 1650
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 HisAaPryGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 543
DB      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-09-949-016-6426
; Sequence 6426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6426
; LENGTH: 559
; TYPE: PRN
; ORGANISM: Human
US-09-949-016-6426

Alignment Scores:
Pred. No.: 4,44e-112 Length: 559
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 38.80% Indels: 24
  
```

```

DB: 4 Gaps: 8
US-10-023-515-3 (1-1746) x US-09-949-016-6426 (1-559)
QY 91 GGGCTTGTGCTGAAGGGCCACAGAGAACACAGGCTGGATTCAGGGCAAG 150
Db 26 G1ylnAaPSeRlASeRPro1LeaG1Thr1ThN1sThG1yAla1Leu1d1ySeR1eu 45
QY 151 GTCACTGTGCTGGAGAGCCCTGCTGCTGAAAGTGTTCAGAGTCCCTTTCTGCT 210
Db 46 ValN1sVal1ySg1yAla1Aen1a1aG1yAl1G1n1Thr1Phe1euG1y1LeP1roPhe1a1y 65
QY 211 CCCCCGTGGAGTCCCTGGATTTAGAACCCGAGCCCTGATCCGCTGGAGTAATGT 270
Db 66 ProP1ro1euG1yPro1eu1d1yPhe1a1P1roP1roG1uP1roG1uSeR1eR1yAl 85
QY 271 CGAGAGCCACTCTTACCTTAATTTGCTCCAGAAC-----TCAGAG 315
Db 86 ArgAaRG1yTh1ThN1sP1ro1A1e1c1ySe1uG1nAaP1e1uTh1A1A1G1uSeR1u 105
QY 316 TGGCTGTCTTAGATCAACATGCTCAAGGTGATTAACCCGAAATGCGAGTGCAGA 375
Db 106 Phe1eu-----SeR1n1Phe1a1n1e1c1Th1r1P1roSe1a1PSe1e1SeR1u 121
QY 376 GACTGCTCTACCTGACATCTATGCGCTGCCACGCGATACAGGCTCCAACTCC 435
Db 122 AaP1ySe1u1yTh1e1SeR1e1yTh1r1P1ro1A1n1sSeR1n1sG1u1ySe1a1n1e1uP1ro 141
QY 436 GTCTTGCTGTGCTCCAGAGGTGCTTCAAGACTGCTCAGCTCTTTGATGG 495
Db 142 ValMe1sVal1Th1P1e1n1sG1yAla1e1uVal1Phe1G1Me1a1Se1e1u1yTh1r1P1e1y 161
QY 496 TCCGCGCTGCTGCTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 555
Db 162 SeR1e1e1u1A1a1e1uG1nA1e1n1A1a1Val1A1e1e1G1n1Th1r1G1e1uG1yAl 181
QY 556 TTTGGTTTCTTACACATGGAGTACAGATGCTCCGCGAATGGGCTTCAAGGAC 615
Db 182 LeuG1yPhe1Phe1SeR1Th1G1yAaP1yN1s1Th1A1Th1G1yAe1n1Th1r1G1y1e1uP1ro1 201
QY 616 GTGCTGTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Db 202 Val1A1a1e1u1A1e1u1G1Th1r1G1n1A1e1n1A1e1a1n1sPhe1G1y1AaP1ro1A1y 221
QY 676 GTGACCATCTTTGGAGTCCGCGGAGCCATTAAGTTTCTAGCTTATCTGCTGCT 735
Db 222 ValTh1r1e1Phe1G1y1e1uSeR1A1e1G1y1Th1r1SeR1A1SeR1e1uVal1A1SeR1o 241
QY 736 ATGCGCAAGGCTTATTCACAAAGCCATGAGAGTGGAGTGGGCAATCCCTTAC 795
Db 242 11e1SeR1n1G1y1e1uPhe1n1sG1yAla11e1Me1c1G1uSeR1G1yAla1A1e1uP1roG1y 261
QY 796 CTGAGGCGCCATGATTAAGAGAGTGAAGCTG-----CAGTGTGTCACATTC 849
Db 262 Leu1e1a1-----Se1SeR1A1a1P1Val11e1SeR1Th1Val1A1A1a1n1e1u 277
QY 850 TGTGTAACAATGGGCTGAGAGCTGAGGCTGCTGAGTGGCTGAGAGCAAACTCC 909
Db 278 SeR1A1c1yAaRG1n1Val1AaPSeR1u1A1e1uVal1G1y1e1uA1yG1y1e1ySeR1y 297
QY 910 AAGAGCTGCTGACCTCAGCCAGAAAACAAGTCTTCACTCGAGTGGTGAATGCT 969
Db 298 GluG1u11e1e1u1A1e1A1e1n1yP1roPhe1yMe1c11e1P1roG1yAl1Val1AaP1G1yAl 317
QY 970 TTTCTTCTATAGCTCTGATCTATTTGTCTAGAAAGATTAAGCAATTCCTCC 1029
Db 318 Phe1euP1roA1yG1h1sP1roG1n1G1u1e1u1A1SeR1A1a1P1he1n1P1roVal1P1roSeR 337
QY 1030 ATCATCGAGCTCAATACCAAGAGTGTGCTTCTGCTGCT-----ATGAAG 1077
Db 338 11e1Val1G1yAla1AaP1e1n1G1uPhe1G1yTh1r1e1u11e1P1ro1y1Sval1Me1a1R1G11e1yR 357
QY 1078 GAGGCTCTGAGATCTCTAGTGGCTCCAACAAGTCCCTTGCCCTTCATCTGATACAAAAC 1137

```

```

Db 358 AaPTh1r1G1n1ySg1uMe1AaPArgG1u1A1SeR1n1A1A1e1uG1n1yMe1e1uTh1r 377
QY 1138 ATCTGCACATCCCGGCTCAGTATTTGACCTTGGGCTAATGAATCTTCCAGACAAG 1197
Db 378 Leu1e1uMe1e1uP1roP1roTh1r1Phe1yAaP1e1uA1yG1uG1y11e1G1yAaP1e1n 397
QY 1198 CACTCCCTGAGTGAATCCGAGACAGTCTTGGACTTGGAGATGATGTTCTTGTG 1257
Db 398 G1yAaP1roG1n1Th1r1e1uG1n1A1aG1n1Phe1n1G1uMe1c1A1a1PSe1e1Phe1a1 417
QY 1258 GTCCCTGACTGTATCAGCTGATATATCAGAGATGCTGCTGACCTGTCTTCTAT 1317
Db 418 11eP1ro1A1e1uG1n1Val1A1n1sPhe-----G1nc1ySe1SeR1A1a1P1roVal1y1Phe1yR 436
QY 1318 GAGTTTGGGACCGGCTCAGTGCCTTTGAACACAGAACCGGCTTTGTCAAGCCGAC 1377
Db 437 GluPhe1G1n1sG1n1P1roSe1Th1r1e1u1yAaP1e1n11e1a1yP1roP1ro1n1sMe1y1A1a1aP 456
QY 1378 CAGCTGATGAAGTCCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
Db 457 H1sG1yAaP1e1uP1roPhe1Val1Phe---ArgSeR1Phe1e1G1y1Ae1n1Th1r11e1y 475
QY 1438 TTGAAAGAGCCAGAGAGAGAGATTACTGAGCCGGAAGATGAATTAAGTGGCT 1497
Db 476 Phe-----Th1r1G1uG1uG1uG1uG1n1e1uSeR1Arg1ySe1e1Me1y1yTh1r1A1a 492
QY 1498 ACCTTGTGCTGACCGGGAATCTTAATGGAGACACTGTCTGCTGCGGCTTAATAT 1557
Db 493 AaPhe1A1a1yAaP1e1n1G1yAaP1roA1e1n1G1y1e1u1yP1ro1h1sTh1r1P1ro1e1uP1ro 512
QY 1558 CTGACTGAGCAGTACTCTCAGCTGAGCTTGAACATGAGAGCTCGGAGAGACTCAAGAA 1617
Db 513 G1n1G1uG1n1yTh1r1e1uG1n1e1uA1e1n1e1uG1n1P1ro1h1A1a1G1yAla1e1u1yA1a 532
QY 1618 CCGCGGCTGATTTTGGACCAACCATCCCT 1650
Db 533 H1sA1rG1e1uG1n1Phe1Th1r1y1yS1A1e1uP1ro 543

RESULT 4
US-09-949-016-9670
; Sequence 9670, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1.001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PASTSEQ for Windows Version 4.0
; SEQ ID NO 9670
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9670

Alignment Scores:
Pred. No.: 4.53e-112 Length: 577
Percent: 1230.00 Matches: 248
Score Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 38.80% Indels: 24
DB: 4 Gaps: 8
US-10-023-515-3 (1-1746) x US-09-949-016-9670 (1-577)

```

```

OY 91 GGGCTTGTGGAAGGGCCACAGAGAACACAGAGCTGGATGTTCAGGCAAGCA 150
DB 44 GAGTAApSerAlaSerPro1LeuGThrThn1stnRglnValLeuG1SerLeu 63
OY 151 GTCACTGTGTGGGAAGCCCTGTGCTGTGAACSTGTTCCTGGAAGTCCCTTGTCTGT 210
DB 64 ValHisValIyG1AlaHisAlaG1ValGlnThrPheLeuG1Y1LeProHehAlaIy 83
OY 211 CCCCCTGGGAGATCCCTGGATTTACGAACCCGACGCTGCATGCGCCCTGGATTAATTG 270
DB 84 ProProLeuG1YrProLeuAlrPheAlaProProGluProGluSerTrpSerG1YVal 103
OY 271 CGAGAAGCCACCTCTACCTAATTGTGTGCTCCAGAAC-----TCAGAG 315
DB 104 ArgHaeG1YrThnHisProAlaMetCysLeuGlnHaeRLeuThAlaValG1SerGln 123
OY 316 TGGCTGTCTTAAATCAACAATGCTCAAGGTGCATTAACCCGAAATTGGAGTGTCAAGA 375
DB 124 PheLeu-----SerGlnPheAlaMetThrPheProSerAlaSerMetSerGln 139
OY 376 GACGTGCTTACCTGAACATCTATGCGCTGCGCCACGCGGATACAGCTCCAAAGCTCCG 435
DB 140 AspCysLeuYrLeuSerTr1eYrThrProAlaHisSerHisG1GlySerAlaLeuPro 159
OY 436 GTCTTGTGTGGTCCCAAGAGGTGCTTCAAGAGCTGCTCAAGCTCTTGAATGGG 495
DB 160 ValMetValTrp1LeHisG1YAlaLeuVal1PheG1YMetAlaSerLeuYrAspG1Y 179
OY 496 TCCGCGCTGCTGCTATAGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
DB 180 SerMetLeuAlaAlaLeuGlnAlaValValVal1Le1eGlnTrYrArgLeuG1YAla 199
OY 556 TTTGGTCTTCAACAATGGAGTCAAGATGCTCCGAGGAACTGGGCGCTTCAAGACAG 615
DB 200 LeuG1YrPhePheSerThrG1YAspYrHisAlaTrpG1YAsnTrpG1YrYrLeuAspGln 219
OY 616 GTGGCTGTCTGTCTGCTGCTCCAGAAACAATCGAGTCTTCGAGGGGAGCCCGAGCTCT 675
DB 220 ValAlaAlaLeuAlaG1YrValGlnGlnAlaHis1LeAlaHisPheG1YAlaAsnProAspArg 239
OY 676 GTGACATCTTTGGGAGAGTCCGCGGAGACCAATAGTGTTCATGCTTAACTGTCTGCC 735
DB 240 ValThrTr1ePheG1YLeuSerAlaG1YLeG1YrHisSerValSerSerLeuValValSerPro 259
OY 736 ATGGCCAAAGCTTATTCACAAAGCCATCATGAGAGTGGGCTGACATCCCTTAC 795
DB 260 I1eSerGlnG1YLeuPheHisG1YAlaTr1eMetGlnSerG1YValAlaLeuLeuProG1Y 279
OY 796 CTGAGAGCCCATGATTAAGAGAGTGAAGCTG-----CAGTGTGTTGCAATTC 849
DB 280 Leu1LeAla-----SerSerAlaAspVal1I1eSerThrValValAlaHisAlaLeu 295
OY 850 TGTGGTAAACAATGCGTCAAGCTCTGAGGCGCTGCTGAGGTGCTGAGGACAAACCTTCC 909
DB 296 SerAlaCysAspGlnValAspSerGlnAlaLeuValG1YrCysLeuAlaG1YrYrSerYr 315
OY 910 AAGAGCTGTGAGCCCTCAAGCCAAAGAAAGTCTTTCATCGAGTGTGATGCTGTCT 969
DB 316 GluGlnTr1eLeuAla1eAlaHisAlaAsnYrProPheYrMetTr1eProG1YValValAspG1YVal 335
OY 970 TTTCTTTCTTAATGAGCTCTAATCTATTTGTCTCAAGAAAGCTTAAAGCAATTCCTTCC 1029
DB 336 PheLeuProAlaGln1sPheGlnG1YrLeuLeuAlaSerAlaAspPheGlnProValAlaProSer 355
OY 1030 ATCATCGAGTCAATTAACAACAGAGTGTGCTTCTGCTGCTCT-----ATGAAG 1077
DB 356 I1eValG1YrAlaHisAlaHisAlaHisGln1YrPheG1YrPheLeu1eProYrValMetArgTr1eYr 375
OY 1078 GAGGCTCTGAGATCTCTCAAGTGTGCTCAACAAGATCCCTTGCCTTCATCTGATTAACAAAC 1137
DB 376 AspThrGlnTr1eYrG1YrMetAspArgG1YrAlaSerGlnAla1eLeuGlnTr1eYrMetLeuThr 395

```

```

OY 1138 ATCTGCATATCCGCTCAAGTATTTGCAACCTTGTGGTAAATGAATTAATTGCATGACAAG 1197
DB 396 LeuLeuMetLeuProProThrPheG1YrAspLeuLeuAlaGlnGlnTr1eG1YrAspAla 415
OY 1198 CACTCCCTGCATGAATAATCCAGACAGCTCTTCTGCACTTGTGGAGATGTGTTCTTTGG 1257
DB 416 GlYAspProGlnThrLeuGlnAlaGlnPheGlnGlnMetMetAlaAspSerMetPheVal 435
OY 1258 GTCCCTGACATGATACAGCTGCATATCAACAAGATGCGGGTGCACCTGTCTAATCTTAT 1317
DB 436 I1eProAlaLeuGlnValAlaHisPhe---GlnCysSerAlaGalaProValYrTrPheTrYr 454
OY 1318 GAGTTCCGACACCGGCTCAAGTGTCTTGAAGACAGAACCGGCTTTTGTCAAGCCGAC 1377
DB 455 GlnPheGlnHisGlnProSerTr1ePheYrAsn1LeaGProProHis1MetCysAlaAsp 474
OY 1378 CACGCTGATGAATGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1437
DB 475 HisG1YrAspGlnLeuProPheValPhe---ArgSerPhePheG1YrG1YrAsnTrYrTr1eYr 493
OY 1438 TTGCAAGAGCCACGAGAGAGAGAAATTACTGAGCCGGAAGATGATGAATACTGGGCT 1497
DB 494 Phe-----ThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 510
OY 1498 ACTTTGTCTGAAACCGGGAATCTTATGGAACGACCTGTCTGTGGCCAGCTTATAT 1557
DB 511 AsnPheAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 530
OY 1558 CTGACTGACAGTACTCTCAAGCTGGAATTTGAACATGAGCCTTGGACAGAGACTCAAGAA 1617
DB 531 GlnGlnGlnGlnTr1eYrLeuGlnGlnLeuAsnGlnProAlaValG1YrArgAlaLeuYrAla 550
OY 1618 CCGCGGTGAGATTTTGTGACACAGCAACCATCCG 1650
DB 551 HisArgLeuGlnPheTr1eYrYrAlaLeuPro 561

```

RESULT 5

```

US-09-595-682B-21
/ Sequence 21, Application US/09595682B
/ Patent No. 6800483
/ GENERAL INFORMATION:
/ APPLICANT: Danke, Mary K.
/ APPLICANT: Poter, Philip M.
/ APPLICANT: Houghton, Peter J.
/ TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
/ FILE OF INVENTION: Tumor Cells
/ FILE REFERENCE: SJ-0005
/ CURRENT APPLICATION NUMBER: US/09/595,682B
/ PRIOR FILING DATE: 2000-01-16
/ PRIOR FILING DATE: 60/075,258
/ PRIOR FILING DATE: 1998-02-19
/ PRIOR APPLICATION NUMBER: PCT/US99/03171
/ NUMBER OF SEQ ID NOS: 30
/ SOFTWARE: PatentIn Ver. 2.0
/ SBO ID NO: 21
/ LENGTH: 565
/ TYPE: PRT
/ ORGANISM: Oryctolagus cuniculus
US-09-595-682B-21

```

Alignment Scores:

```

Pred. No.: 3,05e-104 Length: 565
Score: 1150.50 Matches: 252
Percent Similarity: 56.97% Conservative: 75
Best Local Similarity: 43.90% Mismatches: 180
Query Match: 36.29% Indels: 67
DB: 4 Gaps: 13

```

US-10-023-515-3 (1-1746) x US-09-595-682B-21 (1-565)

```

OY 34 TGGTGTCTTTTCTGATTTCTCCAGCCCTGTGGGACACAGACAGTGGGGAATAAAGCTGGG 93
DB 34 TGGTGTCTTTTCTGATTTCTCCAGCCCTGTGGGACACAGACAGTGGGGAATAAAGCTGGG 93

```



```

QY 34 TGGTCCTTTTCCGATTCCTCCAGCCCTGTGGAGCACAAGAGTGGGAAAACTGGG 93
DB 9 Trrleuphe-----Lenuaiaacythraltipolyhis----- 20
QY 94 CTTTGTGCTGAAGGCCACAGAGAACACAGGCTGGATTCAGAGGCAAGCAATGC 153
DB 21 Proseralapro---ProvalvalapThrVallysglylyvalleuolylusPheVal 39
QY 154 ACTGTGCTGGAGAGCCGTGTGCTGGAAACGTTCTCTGGAGTCCCTTGTCTGCTGCC 213
DB 40 SerleugluglyPhealaglnProvalalavalPheleuglyvalProPhealalyapPro 59
QY 214 CCGCTGGATCCCTGCGATTAACGAACCCGAGCCCTGCATCCCTGGAGTAATCTGCGA 273
DB 60 ProleuglyserleuamigPhealaproProglinProalagluserTripsenhiSallys 79
QY 274 GAAGCCACTCTCACTTAATTTGTCCTCCAGAACTCAGAGTGCCTGCTTACATCAA 333
DB 80 AenThrThrserlyrProPmetCysserglnaPala-----Valsergly 95
QY 334 CACATGCTC-----AAGTGCAATTAACCG---AAATTCGAGTGC 369
DB 96 HisMetleuSerlyleuPheThranaarglysgluasnlleProleuylaphe----- 113
QY 370 TCAGAGACTGCTCTACCTGAACATCTATGCGCTGCCCAAGCCGATACAGGCTCAAG 429
DB 114 SergluapPcybleuThryleuasnilleThrProhlaPheleuThrlyaraglyarag 133
QY 430 CTCCCGCTTGTGTGTGCTCCAGAGAGTGCCTTCAAGACTGCTCAAGCTTCATCTTT 489
DB 134 LeuProvalmetValTrrp1ehisglylyleuMetValglylyalaserThrlyr 153
QY 490 GATGGCTCCGCTGCTGCTATAGAGACGTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 549
DB 154 AapglyleuAlaleuSerliahisglunsnValvalValThrilegintyrargleu 173
QY 550 GGAATATTTGGTTTCTTCAACAGATGAGATGCTCCGAGGAACTGGGCTTCAAG 609
DB 174 GlylletrrpPglyPhePheSerThrglyapglunhisSerArgglyAenThrpllyhisleu 193
QY 610 GACCAAGTGTGCTGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 669
DB 194 AapgluValargAlaleuargTrrpValglnaPasnilleAlaAenPheolylglylaspPro 213
QY 670 AGCTTGTGACCATCTTTTGGCGAGTCCCGGGAGCCATPAAGTGTCTTACTGCTTACTG 729
DB 214 GlySerValThrilePheolylglunSerliahisglunSerValserlileuLeuLeu 233
QY 730 TCTCCATGGCCAAAGGCTTATTCCAAAGCCATCATGAGATGGGGTGGCCATCATC 789
DB 234 SerProleuThrlyasleuPhehishargAlaileSergluSerglyAlaAlaleuLeu 253
QY 790 CTTTACCTGAGGCCCATGATTAAGAGAGTGAAGC-----CTGCAAGTGTGTGCA 843
DB 254 SerSerleu-----PhearglysaenThrlySerleuAlagluylusleuAla 269
QY 844 CATTTCTGTGTAAATAGCGTCAAGACTGTGAGGCCCTGCTGAGTGTGCTGAGCAAAA 903
DB 270 IleglualaglyCyalyethrThrTrhseralavalMetValhiscySeuarglnlys 289
QY 904 CCGTCCAGAGAGTGTGCTG-----ACCTCAGCCAGCAAA----- 936
DB 290 ThrgluglugluMetleuValThrleuLyMetLyPheMetAlaleuapleuVal 309
QY 937 -----ACAAAGTCTTTCACCTGACGAGTGTGTGAGTGTGCTTCTTCTTCT 978
DB 310 GlyapProllysgluasnthralapheleuThrThrValIleapglyValleuLeuPro 329
QY 979 AATGAGCTCTGATATCTATGCTCAAGAAAGCAATTAAGCAATTCCTTCATCATCGA 1038
DB 330 LyvalapProlagluileuAlagluLylyrlyasMetleuProlyrmetValgly 349
QY 1039 GTCAATTAACAGAGTGTGTGCTTCTGCTGCTATG----- 1074

```

```

DB 350 IleaenglnglnPheglyTrrplellePmetGlnMetleuGlyTyrProleuSer 369
QY 1075 -----AAGAGGCTCTCTGATGATCTTCAGAGGCTCCAAAGTCCCTT 1116
DB 370 GluglyLylyleuapglunlyethrAlathrgluLeuThrlylyser----- 385
QY 1117 GCCCTCATCTGATCAAAACATCTCGACATCCCGCTCAGATTAATTCAGCACTGTGCT 1176
DB 386 -----TyrProleValaenValserlysgluLeuThrProvalAlaThr 400
QY 1177 AATGAATCTTCATGACAGACATCCCTGACTGAATCCGAGACATCTTGAACCTTG 1236
DB 401 GluLyTyrrleuglyGlyThrAspAspProvallylylyAspAspPheleuapMet 420
QY 1237 CTGAGAGATGCTTCTTGTGTGCTCCCTGACATGATCAACAGCTCGATATCAGAGATGCT 1296
DB 421 LeuAlaPheleuPheolylvalProserValaenValAlaArgHisargAspAla 440
QY 1297 GGTGACCTGTCTACTTCTATGATGTTTGGCAACCGGCTCAGTGTCTTGAAGACGAAG 1356
DB 441 GlyAlaProThrlyrMetlyrGluTyrrarglyrargProserPheSerSerapheArg 460
QY 1357 CCGGCTTTTGTCAAAAGCCGACCAAGCTGATGAAGTCCGCTTGTGTGTGTGTGTGTGT 1416
DB 461 ProllyThrValIleGlyAspHisglunlePheSerValleuolylalaproPhe 480
QY 1417 CTGAAGGGGACATTTGATGCTTCAAGAGCCACAGAGAGAGAGAACTTACTGAGCCGG 1476
DB 481 LeuLyly-----GluGlyAlaThrTrgluglnGluIlelyleuSerlyrs 494
QY 1477 AAGATGATGAATACTAGGCTACTTGTGCTCGAACCGGGAATCTTAATGGAACGACTG 1536
DB 495 MetValMetlyrlyTrrpAlaAsnPheAlaArgaenglyAsnProAsnolylglunGlyleu 514
QY 1537 TCTGTGCGCCAGCTTATATCTGAATGAGACGATGCTTCAGCTGCACTTGAACATGAGC 1596
DB 515 ProglinThrProalalyAspTyrrlysgluglyTyrrleuGlnIleGlyAlaThrTrgln 534
QY 1597 CTCGACAGAGCTCAAAAGACCGGGGTGAGATTTTGGACC 1638
DB 535 AlaAlaGlnlyleuLyAspLysgluValAlaPheThr 548

RESULT 7
US-09-595-682B-26
; Sequence 26, Application US/09595682B
; Patent No. 6800463
; GENERAL INFORMATION:
; APPLICANT: Danke, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; PRIOR FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 543
; TYPE: PR
; ORGANISM: Oryctolagus cuniculus
US-09-595-682B-26

Alignment Scores:
Pred. No.: 2,796-102 Length: 543
Score: 1130.50 Matches: 249
Percent Similarity: 56.84% Conservative: 75
Best Local Similarity: 43.68% Mismatches: 179

```


CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/017,879
 FILING DATE: 17-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Cheryl J. Tubach
 REGISTRATION NUMBER: 38,346
 REFERENCE/DOCKET NUMBER: 70432
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 423-229-6189
 TELEFAX: 423-229-1239
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 584 amino acids
 TYPE: Amino acid
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 US-08-845-295A-2

Alignment Scores:
 Pred. No.: 2,036-98 Length: 584
 Score: 1091.50 Matches: 259
 Percent Similarity: 54.39% Conservative: 63
 Best Local Similarity: 43.75% Mismatches: 191
 Query Match: 34.43% Indels: 79
 Gaps: 17

US-10-023-515-3 (1-1746) x US-08-845-295A-2 (1-584)

```

QY 34 TGGTCTTTTCTGATCTCCAGCCCTGTTGGACACAGACAGTGGGAAATGCG 93
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2 TripleuleuproleuValleuthrSerleuAlaThr-----AlaGly 19
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94 CTTTCTGTAAGGCCACAGAGAACACAGCTGGGATTCAGGCGAAGCATC 153
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 20 GlnProAlaSerProProValValAlaSerThrAlaGlnGlyArgValleuGlySerVal 39
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 154 ACTGG-----CTGGGAAGCCCTGCTGCGAAGCGTCTCCGGAGTCCCTTGT 207
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 40 SerleuGlnGlyLeuAlaPheThrGlnProValAlaValPheLeuGlyValProPheAla 59
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 208 GCTCCCGCGCTGGGATCTCCGATTTACGAACCCGACGCTGCATCCCGCTGGGATAC 267
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 60 LysProProleuGlnSerleuAlaPheAlaProProGlnProAlaGlnProThrSerPhe 79
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 268 TTGGAGAGCCACCTCTACCTTAATTGCTCTCCAGAACTCA-----GAGTGGCTG 321
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 80 ValLysAenThrThrSerLysProPheMetCysGlnAerProValValGlnGlnMet 99
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 CTCCTTAAGT-----CAACATGCTCAAGGTGCATTAACCGAA 360
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 ThrSerAerPhePheThrAerPheThrGlyLysGlnAerGlnPheLeuGlnPhe----- 117
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 TTGGAGAGTGGAGAGACTGCTTACCTGAACATCTATGCGCTGCCACGCGATACA 420
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 118 -----SerGlnAerPheLeuThrLeuAlaValPheThrValValValAla 134
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 GGTTCGAAGCTCCCGCTTGGTGGTCCAGAGAGTGCCTTAAGAAGTGGCTCAGCC 480
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135 ArgGlnAerGlnProValMetValThrLeuAlaGlnGlyGlyLeuValleuGlyAla 154
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 TCCATCTTTGATGGTCCCGCTGGCTGCTATAGAGAC-----GTGCTGTTGGTGC 534
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 ProMetLysAerGlnValValleuAlaValPheThrValValValAla 174
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 535 GTCCAGTACCGGCTAGATATTTGTTCTTCAACATGGGATCAGCATGCTCCGGG 594
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 ILeuGlnThrGlnGlnGlyLysThrPheSerThrGlnAerGlnAerGln 194
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 AACTGGGCTTCAAGACACGAGTGCCTGCTGCTGGGCTCAGAAAGACATGCTTC 654
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 AenThrGlnLysLysAerGlnValAlaAlaLeuThrThrValGlnGlnAerAlaAen 214
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 655 TTCGGTGGGAGCCCGAGCTGTGACCATCTTGGCGAGTCC-----GCGGAGCCATA 708
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 215 PheGlnGlyAerProGlnSerValThrLeuGlnGlnSerPheThrAlaGlnGly 234
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 709 AGTGTCTTACTTTATCTGTCTCCATGGCGAAAGCTTATTCACAAAGCATCATG 768
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 SerValSerValleuValleuSerProleuAlaLysAenPheThrValAlaLeuSer 254
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 769 CAGAGTGGGCGGACCATCCCTTACCTGAGGCGCCATGATTAAGAGTGAAGGAC 828
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 GlnSerGlnValAlaLeuThrValAlaLeuValArgLysAerMetLysAlaAlaLys 274
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 829 CTGCAGTGGTTGACATTTCTGTGTAAACAATGCGCTGACATCTGAGGCGCTGGAG 888
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 275 GlnIleAlaValleuAlaGlyCysLysThrThrThrSerAlaValPheThrPheAlaHis 294
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 889 TCCCTGAGACAAACCTCCAGAGAGTGTGACCTTCAGCCAGAAACAAAGCTTTC 948
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 CysLeuArgGlnLysSerGlnAerGlnLeuAerPheLeuThrLeuLysMetLysPheLeu 314
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 949 ACT-----CGAGTGGTGTAT 963
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 ThrLeuAerPheThrGlyAerGlnArgGlnAerThrSerProPheLeuProThrValAlaLys 334
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 964 GGTGCTTTCTTTCTAATGAGGCTTGAATCTATTGCTCAGAAA-----GCATTTAAA 1017
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 335 GlnValleuLeuProLysMetProGlnGlnLysLeuAlaGlnLysAerPheThrPheAla 354
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1018 GCATTTCTTCATCATGAGAGTCAATTAACAGAGTGGCTTCTGCTGCT----- 1071
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 ThrValProGlyThrLeuAlaGlyLysAerGlnGlnPheGlnThrLeuLeuProThrMet 374
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1072 -----ARGAGAGGCTCTGAGATCCCTCACT 1098
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 MetGlyPheProleuSerGlnGlyLysLeuAerGlnLysThrAlaThrSerleuLeuThr 394
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1099 GGTCCCAACAAGTCCCTTGCCCTCATGTGATACAAACATCTGCACATCCCGCTCAG 1158
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 395 LysSer-----TyrProIleAlaAenIle-----ProGln 404
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1159 TATTGACCTTGGCTAATGATATCTTCATGCAAGAC-----TCC 1203
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 405 GlnLeuThrProValAlaThr-----PheThrAerLysGlyLeuGlyGlyLysAer 422
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1204 CTGACTGAATCCGAGACAGCTTCTGACTTGTGAGAGATGTTCTTGTGGCTCT 1263
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 ProValLysLysAerPheLeuAerPheLeuMetGlyAerValValPheGlyValPro 442
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1264 GCATGTATCAGAGCTGCATATCAGAGATGCTGGGACCTGTCTACTTATGAGTT 1323
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 SerValThrValAlaArgGlnIleArgAerAlaGlnAlaProThrTyrMetLysGlnPhe 462
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1324 CCGCACCGGCTCAGTGTGTAAGAC-----ACGAACCGGCTTGTCAAGCCGAC 1377
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463 GlnTyrArgProSerPheSerSerAerLysPheThrLysProLysThrValIleGlyAer 482
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1378 CAGCTGATGAAGTCCGCTTGTGCTGGTGGCTCTTCAAGAGGAGCAATTGTATG 1437
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 483 HisGlyAerGlnIlePheSerValPheGlyPheProleuLeuLysGlyAer----- 499
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1438 TTGGAAGGACGAGAGAGAGAAAGTTACTGAGCCGAGATGATGAATATAGTGGCT 1497
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 500 -----AlaProGlnGlnGlnValSerleuSerLysThrValMetLysPheThrAla 516
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1498 AACTTGTCTGAAACCGGAGATCTAATGGAACGACTGTCTGTGGCA-----GCT 1551
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 517 AenPheAlaLysSerGlyAerProAenGlnGlnGlyLeuProThrLysProPheThrMet 536
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1552 TATTAATCTGACTGAGACAGTACTCAGCTGAGACTTGAACATGAGCTCGAGACAGCTC 1611
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 537 TyrAerGlnGlnGlnGlyLysLeuGlnIleGlyValAenThrGlnAlaLysAerGln 556
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1612 AAAGAACCGGCGGATTTTGGACAGACCATTC 1647
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Db	295	CysLeuAlaGlnLysSerGlnAspGluLeuLeuAspLeuThrLeuLysMetLysPheLeu	314
Qy	949	ACT-----	CGAGTGGTGAAT 963
Db	315	ThrLeuAspPheHLeGlyAspGlnArgGluSerHLeProPheLeuProThrValAlaAsp	334
Qy	964	GGTGGTCTCTTCTTCCTAATGAGCGCTTAAGATCTATATGTCTCAGAA-----GCATTATAA 1017	
Db	335	GlyValLeuLeuProLysMetProGluGlnLeuAlaGluLysAspPheThrPheAsn	354
Qy	1018	GCAATTCCTTCGCATCATCGAGAGTCAGATAACACAGATGGCTTCCTGCTGCT----- 1071	
Db	355	ThrValProTyrLeuAlaGlyLeuAlaGlnLysGlnLysPheGlyTyrPheLeuProThrMet	374
Qy	1072	-----ATGAAGAGGCTCTGAGATCCTCAGT 1098	
Db	375	MetGlyPheProLeuSerGlnGlyLysLeuAspGlnLysThrAlaThrSerLeuLeuTyrP	394
Qy	1099	GGCTTCGACAAAGTCCCTTCCTCCATCTGATACAAACATCTCTGACATCCCGCTCAG 1158	
Db	395	LysSer-----TyrProLeuAlaAsnIle-----ProLys 404	
Qy	1159	TATTGGACCTTGGCTAATGAACTACTCCATGACAGACAC-----TCC 1203	
Db	405	GluLeuThrProValAlaThr-----PheThrAspLysTyrLeuGlyGlyThrAspAsp	422
Qy	1204	CTGACTGAAATCCGAGACAGTCTTCGTGACCTTGTGACATGTCCTTGTTCGTGCTCCT 1263	
Db	423	ProValLysLysLysAspLeuPheLeuAspLeuMetGlyAspValAlaPheGlyValPro	442
Qy	1264	GCACGTGATCAGACGCTCGATATCACAGAAATGCTGGTGCACCTGTACTTCTATGAGTTT 1323	
Db	443	SerValThrValAlaArgGlnHLeArgAspAlaGlyAlaProThrTyrMetCylLysPhe	462
Qy	1324	CGGACCGGGCCAGTGCCTTGGAAAC-----ACGAACCGGGCTTTGTGTAAGCGGAC 1377	
Db	463	GlnTyrArgProSerPheSerSerAspLysPheThrLysProLysThrValIleGlyAsp	482
Qy	1378	CACGCTGATGAAAGTCCGCTTTGTGTTCGATGGTGCCTTCGAAAGGGAGCATTTGTAAG 1437	
Db	483	HLeGlyAspGlnLysPheSerValPheGlyPheProLeuLeuLysGlyAsp----- 499	
Qy	1438	TTCGAAGAGCCACGAGAGAGAGAAAGTTACTGAGCCGGAAGATGATAAATCTGGGCT 1497	
Db	500	-----AlaProGlnGlnGlnValLysSerLeuSerLysThrValMetLysPheThrAla 516	
Qy	1498	ACCTTTGTGGAACCGGGGAATCTTAATGGGAACGACCTGTCTGTGGCC-----GCT 1551	
Db	517	AsnPheAlaArgSerGlyAsnProAsnLysGlyLeuProHLeTyrProPheThrMet	536
Qy	1552	TATATCTGACTGAGCAGTACTCTCAGCTGAGCTTGAAACATGAGCGCTTCGACAGAGACTC 1611	
Db	537	TyrAspGlnGlnGlyLysTyrLeuGlnIleGlyValAsnThrGlnAlaLysAlaArgLeu	556
Qy	1612	AAAGAACCGCGGTGATTTTGGACGACGACATC 1647	
Db	557	LysGlyGlnGlnValAlaPheThrAsnAspLeuLeu 568	

RESULT 11
 US-09-150-515-2
 Sequence 2, Application US/09150515
 Patent No. 6271006
 GENERAL INFORMATION:
 APPLICANT: Hubbs, John C.
 TITLE OF INVENTION: Enzymatic Process for the Manufacture of
 TITLE OF INVENTION: Acoodic Acid, 2-Keto-L-Gulonic Acid, and Esters of
 TITLE OF INVENTION: 2-Keto-L-Gulonic Acid
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESS: Eastman Chemical Company
 STREET: P.O. Box 511
 CITY: Kingsport

D	155	ProMetcylrAaBpGlnValValLeuAlaIahIbGluuSenPheThrValValValAla	174
Q	535	GTCCAGTACCCGGCTAGAGATATTGGTTCTTCAACACATGGAGTCACATCCGGG	594
D	175	IlleGlnTyraBleuGlyIletrpGlyPhePheSerThrGlyAaBpGluHiseAaBpGly	194
Q	595	AACTGGCCCTTCAAGAACACAGTGGCTGCTGTCTCCGGGTCCAGAGAACATCGAGTTC	654
D	195	AaHtrpGlnHiseLeuAaBpGlnValAlaIaLeuHistrValGlnGluAaSnIAlaAaH	214
Q	655	TTGGTGGGGAGCCCGAGCTGTCAGCACTTTGGCCAGTCC-----GGGGAGCCATA	708
D	215	PheGlyGlyAaBpProGlySerValThrIlePheGlyGlySerPheThrAlaGlyGly	234
Q	709	AGTGTTCCTACTCTTATCTGCTCTCCAGTCCCAAGGCTATTTCACAAAGCATCATG	768
D	235	SerValSerValLeuValLeuSerProLeuAlaIaAaSenLeuPheHiseAaGlaIleSer	254
Q	769	GAGAGTGGGGTGGCCCATCATCCCTTAACTGAGGCCATGATTATGAGAAAGTGAAGAC	828
D	255	GluSerGlyValAlaIaLeuThrValAlaLeuValArgIyAaMetcylValAlaIaIy	274
Q	829	CTGCAGGTCGTGCACATTTCTGTGTGAACAATGCTGACATCTGAGCCCTCTAGG	888
D	275	GlnIleAlaValLeuAlaGlyCysIySerThrThrThrSerAlaValPheThrPheValHise	294
Q	889	TGCCTTAGAGCAAAACCTCCAGAGAGCTGTGCACCTCCAGCCAGCAAAACAAGTCTTC	948
D	295	CysLeuAaBpGlnIySerGlnAaBpGluLeuLeuAaBpLeuThrLeuIyMetcylPheLeu	314
Q	949	ACT-----CGAGTGGTGGAT	963
D	315	ThrLeuAaBpPheHiseGlyAaBpGlnAaBpGluSerHiseProPheLeuProThrValValAaBp	334
Q	964	GGTGTCTTCTTCTTAATGAGCCTCTAGATCTATTGTCTCAGAA-----GCATTATAA	1017
D	335	GlyValLeuLeuProIyAaMetProGlnIleLeuAlaGlyIyAaBpPheThrPheAaH	354
Q	1018	GCAATTCCTCCATCATCTCGAGAGTCAATAACACAGAGTGGCTCTCTGCTGCT	1071
D	355	ThrValProTyrlIeValGlyIleAaSnIySerGlnIyPheGlyTyrlPheLeuProThrMet	374
Q	1072	-----ATGAAGAGCTCTGAGATCTCACT	1098
D	375	MetGlyPheProLeuSerGlnGlyIyLeuAaBpGlnIySerThrAlaThrSerLeuTrp	394
Q	1099	GGCTTCACAAAGTCCCTTGGCCCTCATGTATACAAACATCTGACATCCCGCTCAG	1158
D	395	LysSer-----TyrProIleAlaAaSnIe-----ProGlu	404
Q	1159	TATTGGACCTTGGGCTAATGAAATACTTCATGACAGAC-----TCC	1203
D	405	GluLeuThrProValAlaThr-----PheThrAaBpIyTyrlLeuGlyGlyThrAaBp	422
Q	1204	CTGACTGAATCCGAGACAGTCTTCTGCACTGTCTGGAGATGTTCTTTGTGGTCCCT	1263
D	423	ProValIySlyIyAaBpLeuPheLeuAaBpLeuMetGlyAaBpValValPheGlyValPro	442
Q	1264	GCACTGAATCAGCTTCGATATACACAGAGATGCTGTGCACTGTCTACTTCTAAGAGTTT	1323
D	443	SerValThrValAlaArgGlnHiseAaBpAlaGlyAlaProHTrIyMetcylIyPhe	462
Q	1324	GGGACCGGCGCTCAGTCTTGAAGAC-----ACGAAGCCGGCTTTGTCAAGCCGAC	1377
D	463	GlnTyraBpProSerPheSerSerAaBpIyPheThrIyProIySerThrValIleGlyAaBp	482
Q	1378	CACGCTGATGAAGTCCGTTTGTTTCCGATGGTGCCTTCGAAAGGGGACATTTGTTATG	1437
D	483	HiseGlyAaBpGluIlePheSerValPheGlyPheProLeuLeuIyGlyAaBp-----	499
Q	1438	TTCCAGAGGCCACGAGAGAGCAAGTTACTAGCCCGAAGATGATGAATACTGGGCT	1497

```

Db      500 -----AlaProGluGluGluValSerLeuSerLysThrValMetLysPheThrIle 516
Qy      1498 ACCTTGCTGCAGACCGGAGATCTTAATGGAGAACACTGTCTCTGGGCCA-----GCT 1553
Db      517 AsnPheAlaArgSerGlyAsnProAsnGlyGluGlyLeuProHisLysPheThrMet 536
Qy      1552 TATATATCGATGTCAGACGACTACCTCCAGCTGGACCTTGAAACATGAGCGCTCGACAGAGACTC 1611
Db      537 TyrAspGluGluGluGlyTyrLeuGlnIleGlyValAsnThrGlnAlaIleAlaLysArgLeu 556
Qy      1612 AAGAACCGCGGTGATTTTGGACCGACCACTC 1647
Db      557 LysGlyGluGluValAlaPheThrAsnAspLeuLeu 568

RESULT 12
US-09-264-737-1
; Sequence 1, Application US/09264737A
; Patent No. 6107549
; GENERAL INFORMATION:
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Ruff, Thomas G.
; TITLE OF INVENTION: Engineering Plant Resistance to Pyridines via
; FILE REFERENCE: 38-21(10551) RUS3 Pyridine Tolerance
; CURRENT APPLICATION NUMBER: US/09/264,737A
; CURRENT FILING DATE: 1999-03-09
; EARLIER APPLICATION NUMBER: 60/077,377
; EARLIER FILING DATE: 1998-03-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Rabbit
US-09-264-737-1

Alignment Scores:
Pred. No.: 7,99e-93 Length: 539
Score: 1034.50 Matches: 234
Percent Similarity: 55.05% Conservative: 71
Best Local Similarity: 42.24% Mismatches: 176
Query Match: 32.63% Indels: 73
DB: 3 Gaps: 13

US-10-023-515-3 (1-1746) x US-09-264-737-1 (1-539)
Qy      94 CCTTCGTGGAAGGGCCACAGAGAACCAAGCGTGGATGATCAGGGGACGAATC 153
Db      2 ProSerAlaPro---ProValAlaSerThrValLysGlyValValLeuGlyLysPheVal 20
Qy      154 ACTGTGCTGGAGAACCCCTGTGCTGTGAACTGTTCCTCGAGAGTCCCTTGTGCTGCTCC 213
Db      21 SerLeuGluGlyPheAlaGlnProValAlaValPheLeuGlyValProPheAlaLysPro 40
Qy      214 CCGGTGGATCCCTGCAGATTACGAACCCGACGCTGATGCGCTGGGATAACTTGGCA 273
Db      41 ProLeuGlySerLeuArgPheAlaProProGlnProAlaGluSerTrpSerHisValLys 60
Qy      274 GAAGCCACCTCTACCCCTAATTGTGCTCCAGAACTCAGAGTGCGTCTTATGATCA 333
Db      61 AsnThrThrSerTyrProProMetCysSerSerAla-----ValSerLys 76
Qy      334 CACATGCTC-----AAGTGACATTACCG---AAATTGGAAGTG 369
Db      77 HisMetLeuSerGluLeuPheThrAsnArgLysGluAsnIleProLeuLysPhe----- 94
Qy      370 TCGAAGAAGTCCCTTACCTGAACAATGTATGCGCTGCCACGCGCGATACAGGCTCCAG 429
Db      95 SerGluAspCysLeuTyrLeuAsnIleTyrThrProAlaSerLeuThrLysArgGlyArg 114
Qy      430 CTCGCCGCTTGAGTGTGGTCCCGACGAGGCGCTTCAAGAATGGGTGACGCTCAATCTT 489
Db      115 LeuProValMetValTrpIleHisGlyGlyGlyLeuMetCysValGlyValAsnSerThrTyr 134

```

```

490 GATGGATCGCCCTGCTATGAGACGCTGTTGATGTCAGTACCGGCTA 549
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 AapGlyLeuAlaLeuSerAlaHisGluAenValValValThrIleGlnIlyrAglLeu 154
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
550 GGAATATATTTGTTTCTTCAACAATGGATCAGCATGCTCGGGGAATCGGCTTCAAG 609
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
155 GlyIleGlyIlePheGlyPheAenIleAerGlu-----Leu 166
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
610 GACCAAGTGGCTGCTCTGCTCCGGAAGAAATCGAATCTTCGGTGGGACCCC 669
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 PheLeuValAlaValAasnArgTTPValGlnAasnPheIleAlaasnPheGlyIleAspPro 186
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
670 AGCTCTGACCATCTTTGGGAGTCCGGGGAGCCATTAAGTTTCTGATCTTATACG 729
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
187 GlySerValThrIlePheGlyGlnSerAlaGlyGlnSerValSerIleLeuLeuLeu 206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
730 TCTCCATGCGCAAGGCTTATTCACAAAGCATCATGAGAGTGGGGTGGCCATCATC 789
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
207 SerProLeuThrIlyAasnLeuPheHisArgAlaIleSerGlnSerGlyValAlaLeuLeu 226
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
790 CCTTACCTGGAGGCCCATATATGAAAGAGTGAAGAC-----CTGCAAGTGGTTGCA 843
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
227 SerSerLeu-----PheArgIlyAasnThrIlySerLeuAlaGlyIlySerIleAla 242
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
844 CATTTCTGCTGTAACATGCTGAGACTGTGAGGCCCTGAGGCTGAGGCTGAGACAA 903
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
243 IleGlnAlaGlyCylLeuThrThrSerIleAlaValMetValHisCylLeuArgGlnIly 262
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
904 CCTTCAGAGAGACTCTG-----ACCTCGACCAAGAA----- 936
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
263 ThrGlnGlnGlnLeuMetGlnValThrLeuIlyMetIlyPheMetAlaLeuAspLeuVal 282
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
937 -----ACAAAGCTTTCACTCGAGTGGTGAAGTGGCTTCTTCTTCT 978
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 GlyAspProIlyGlnAasnThrAlaPheLeuThrThrValIleAspGlyValLeuLeuPro 302
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
979 AATGAGCTCTGATCTATTTGCTCAGAAAGCATTTAAAGCAATTCCTTCATCGGA 1038
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 LysAlaProAlaGlnIleIlyrGlnGlnIlySerIlyrAenMetLeuProIlyMetValGly 322
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1039 GTCAATTAACCAAGAGTGGCTCTCTGCTGCTATG----- 1074
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
323 IleAenGlnGlnIlyPheGlyIlyrIleIlePrometGlnMetLeuGlyIlyrProLeuSer 342
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1075 -----AAGAGGCTCTGAGATCTTCAGTGGCTCCAAAGTCCCTT 1116
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 GlnGlyIlyLeuAspGlnIlySerAlaThrGlnLeuLeuTrpIlySer----- 358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1117 GCCCTCATCTGATCAAAACATCTCTGACATCCGCGCTCAGTATTTGCACTTGTGCT 1176
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 -----TyrProIleValAenValSerIlyGlnLeuThrProValAlaThr 373
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1177 AATGAATCTTCATGACACAGCATCTCTGACTGAATTCAGACAGATCTTGTGACTTG 1236
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
374 GlnIlySerIlyLeuGlnIlyThrAspAspProValIlySerIlyAspLeuPheLeuAspMet 393
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1237 CTGGAAGATGCTGTTTGTGTGCTCCCTGACATGATCAGCTCCGATATCAAGAAATCT 1296
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
394 LeuAlaAspLeuLeuPheGlyValIProSerValAenValAlaAArgHisIleArgAspAla 413
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1297 GGTGACACGCTCTACTTCTATGATTCGACACCGGCGCTCAGTGGCTTGAAGACAGAG 1356
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
414 GlyAlaProThrIlyMetIlyrGlnIlyrArgIlyrArgProSerPheSerSerAspMetArg 433
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1357 CCGGCTTTGTCAAAAGCGACACGCTAGTAAGTCCGCTTGTGTGGTGGCTCTTC 1416
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
434 ProIlyThrValIleGlyAspHisGlyAspGlnIlePheSerValLeuGlyValaProPhe 453
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1417 CTGAAGGGGGACATTTGATGTTGCAAGAGCCAGGAGGAGAAATTTCTGAGCCGG 1476
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
454 LeuIlyS-----GlnGlyAlaThrGlnGlnIlyLeuIlyLeuSerIlyS 467

```

```

QY 1477 AAGATGATGAATACTAGGCTTACTTGTCTGAAACCGGGAATCTTAATGGAAACGACTG 1536
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB MetValMetIlySerIlyTrpAlaAsnPheAlaArgAenGlnAasnProAenGlyGlnIlyLeu 487
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1537 TCTGTGGCCAGCTTATATATGACTGACAGATCTTCAGCTGACTGACTTGAACATGAGC 1596
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 488 ProGlnTrpProAlaIlyrAspIlyrIlySerGlnIlyIlyrLeuGlnIleGlyAlaThrThrGln 507
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1597 CTCGACAGACGATCAAGAAACCGGCGGTGATTTTGGACC 1638
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 508 AlaAlaGlnIlyLeuIlyAspIlySerGlnValAlaAenTrpThr 521
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-08-446-100-26
; Sequence 26, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Okana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
; US-08-446-100-26

Alignment Scores:
Pred. No.: 8,37e-84 Length: 454
Score: 942.50 Matches: 200
Percent Similarity: 59.33% Conservative: 64
Best Local Similarity: 44.94% Mismatches: 150
Query Match: 29.73% Indels: 31
DB: 3 Gaps: 6

US-10-023-515-3 (1-1746) x US-08-446-100-26 (1-454)
QY 373 GAAGACGCTCTTACTGGAACATCTATGCGCTCCACGCGCATACAGGCTCCAGCTC 432
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 GlnHisCysLeuIlyrLeuAenIleIlyrThrProAlaAspLeuThrIlySerAasnArgLeu 20
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

```

Db 372 AlsebergIugIugIuIleArgLeuSerIySeWetValMeUtySPheTrpAlAsnPha1a 391
QY 1507 CGAACCGGGAATCTTAATGGAGACGACTGTCTCTGTGGCCAGCTTATTACTGACTGAG 1566
Db 392 ArgAsnGlyAsnProAsnGlyLysGlyLeuProHisTrpProGluTyAsnGlnLysGlu 411
QY 1567 CAGTACCTCCAGCTGAGCTTGACATGAGCCTCGGACAGACACTCAAGAACCGCGGGTG 1622
Db 412 GLyTyLeuGlnGlnIleGlyAlaAsnThrGlnAlaAlaGlnLysLeuTyAspLysGluVal 431
QY 1627 GATTTTGGACGACG 1641
Db 432 AlaPheTrpThrAsn 436

RESULT 14
US-08-446-100-28
; Sequence 28, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-3767
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEITICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
; US-08-446-100-28

Alignment Scores:
Pred. No.: 8.37e-84 Length: 454
Score: 942.50 Matches: 200
Percent Similarity: 59.55% Conservative: 65
Best Local Similarity: 44.94% Mismatches: 149
Query Match: 29.73% Indels: 31
DB: 3 Gaps: 6

US-10-023-515-3 (1-1746) x US-08-446-100-28 (1-454)
373 GAAAGCTGCTCTACTCAATCATCTATGCGCTGCCAAGCGGATATACAGGCTCAAGCTC 432

```

```

Db      1  GlnHisCysLeuTyrLeuAsnIleTyrThrProIleAspLeuThrIleValAsnAsnArgLeu 20
Qy      433  CCCGCTTGGTGTGGTGTCCAGAGAGTGGCTTCAAGACTGGCTCAGACCTCCATCTTGAT 492
Db      21  ProValMetValTyrIleHisGlyIleHisGlyValGlyIleAsnIleAspSerThrTyrAsp 40
Qy      493  GGGTCCGGCTGGCTGGCTATGAGAGCGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 552
Db      41  GlyLeuAlaLeuAlaIleHisGlyIleValIleValIleThrIleGlnTyrArgLeuGly 60
Qy      553  ATATTGGTGTTCCTTCAACCATGAGATCAGATGCTCCGGGAGACTGGGCTTCAAGAC 612
Db      61  IletTyrGlyPhePheSerThrGlyAspGlyHisSerArgGlyAsnTyrPolYHisIleuAsp 80
Qy      613  CAGGTGGCTGCTGTCTGTCTGGGTCCAGAGAACATCGACTTCTTGGTGGGAGACCCAGC 672
Db      81  GlnValAlaIleAsnArgTyrValGlnAspAsnIleAlaSerPheGlyGlyAsnProGly 100
Qy      673  TCTGTGACCATCTTGGGAGAGTCCGGGAGGCCATTAAGTGTTCATAGTCTTATGCTGT 732
Db      101  SerValThrIlePheGlyValSerAlaGlyGlyValSerValSerValIleuValIleuSer 120
Qy      733  CCCATGGCCAAAGGCTTATTCACAAGCCATCATGAGAGTGGGAGTGGCCATCATCCCT 792
Db      121  ProIleuAlaIleAsnLeuPheHisArgAlaIleSerGlyValAlaIleuThrSer 140
Qy      793  TACCTGGAGGCCCATGATTATGAGAAG-----AGTAGGACTCTGACAGTGGTGTGACAT 846
Db      141  ValIleuValIleValIleAspValIleValIleAspValIleAlaIleThrIleAla--- 159
Qy      847  TTCTGTGGTAAACAATGCGCATGACTGTGAGCGCTGTGAGGCTGTGAGCAACAACCC 906
Db      160  -----GlyCysIleThrThrThrSerAlaIleAlaMetValHisCysLeuArgGlnIleThr 177
Qy      907  TCCAGAGAGCTGCTGACCCCTGACCCAGAA----- 936
Db      178  GlnGlnGlnIleuLeuGlnThrThrIleuValIleGlyAsnSerTyrLeuThrThrTyrArg 197
Qy      937  -----ACAAAGCTTTCACCTGACGATGCGTGTGATGAGCTTCTTCTTCTTAT 981
Db      198  GlnThrGlnArgGlnIleThrIleuLeuGlnIleThrIleAspGlyMetLeuLeuLeuVal 217
Qy      982  GAGCCTTGAATCTATGCTCAGAAAGCATTTAAAGCAATCTTCATCATGCGGAGTGC 1041
Db      218  ThrProGlnIleuLeuGlnArgGlnIleAsnPheHisIleThrValProTyrMetValIleGly 237
Qy      1042  AATAACACGAGTGTGGCTTCTGCTGCTATGAAAGAGGCTCTGAGATCCTCAGTGC 1101
Db      238  AsnIleuGlnIleuPheGlyTyrIleuIleProMetGlnIleuMetSerTyrProIleuSerGln 257
Qy      1102  TCC-----AACAACTCCCTGGCTCCATCTGATACAAAC-----ATCCTGCAC 1146
Db      258  GlyIleuLeuAspGlnIleThrIleMetSerLeuLeuGlySerProIleProIleuPheAla 277
Qy      1147  ATCCGCTCAGTATTTGACACTGTGTGCTATGATATCTTCATGACAGACATCCCTG 1206
Db      278  IleAlaIleValIleuIleProGlnAlaThrGlnIleValTyrLeuGlyGlyThrAspAspThr 297
Qy      1307  ACTGAATCCGAGACAGACTTCTGACTTCTGAGATGTGTCTTGTGGTCCCTCA 1266
Db      298  ValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 317
Qy      1267  CTGATCAGACTCCATATACAGAGATGTGTGTGACCTGTCTTACTTCTATGATTTGG 1326
Db      318  ValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 337
Qy      1327  CACCGGCTCAGTGTGTTGAGAGACGAGCGGCTTTGTCAAGCCGACGACGCTGT 1386
Db      338  TyrArgProSerPheSerSerAspIleValIleValIleValIleValIleValIleValIle 357
Qy      1387  GAAATCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1446

```

```

Db      358  GluIleuPheSerValPheGlyAlaProPheLeuIys-----GlnGly 371
Qy      1447  GCCACGGAGGAGAGAACTTCTGAGCGCGGAGAAATGATGAATCTGGGCTACCTTGGCT 1506
Db      372  AlasSerGlnGlnGlnIleArgLeuSerIleMetValIleValIleValIleValIleValIle 391
Qy      1507  CGAACCGGGAATCCATATGAGAACGACCTGTCTGTGGCCAGGCTTATATCTGACTGAG 1566
Db      392  ArgAsnGlyAsnProHisGlyIleValIleValIleValIleValIleValIleValIleVal 411
Qy      1567  CAGTACCTTCAGCTGAGCTTGAACATGAGCTCGGACAGACATCTAAAGAACCGGGGTG 1626
Db      412  GlyTyrLeuGlnIleGlyAlaAsnThrGlnAlaIleGlnIleValIleValIleValIleVal 431
Qy      1627  GATTTTGGACACG 1641
Db      432  AlaPheTyrPheAsn 436

RESULT 15
US-08-446-100-30
; Sequence 30, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: Broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-3767
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
; US-08-446-100-30

Alignment Scores:
Pred. No.: 8,37e-84 Length: 454
Score: 942.50 Matches: 200
Percent Similarity: 59.33% Conservative: 64
Best Local Similarity: 44.94% Mismatches: 150
Query Match: 29.73% Indels: 31
DB: 3 Gaps: 6

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: June 15, 2005, 08:53:20 ; Search time 120.306 Seconds

(without alignments)
11126.674 Million cell updates/sec

Title: US-10-023-515-3

Perfect score: 3170
Sequence: 1 atgcacacagggagcttacttc.....ttctcttctgtcctcttga 1746

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 3420798

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_plus.n2p.model -DEV=x1h
-O=/cgn2_1/USPTO_epool/6664091/runat_14062005_133501_9433/app_query.fasta_1.4238
-DB=Published Applications AA -QMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USR=6664091.@CEN_1_1_295.@runat_14062005_133501_9433
-ICPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US10H_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US10I_PUBCOMB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US10J_PUBCOMB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US10K_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	3079	97.1	581 13	US-10-023-515-2
2	3079	97.1	581 15	US-10-674-636-2
3	3079	97.1	581 16	US-10-757-262-46
4	2915	92.0	575 15	US-10-451-168-91
5	2900	91.5	642 15	US-10-433-256-10
6	2607.5	82.3	581 15	US-10-451-168-92
7	2606	82.2	525 15	US-10-094-749-2275
8	2597.5	81.7	581 15	US-10-114-270-196
9	2588.5	81.9	618 15	US-10-381-898-2
10	2454	77.4	469 15	US-10-104-047-2219
11	1985	62.6	542 16	US-10-233-933A-2
12	1985	62.6	542 16	US-10-233-933A-4
13	1780	56.2	356 15	US-10-451-168-93
14	1230	38.8	549 9	US-09-925-298-689
15	1230	38.8	549 14	US-10-102-806-689
16	1230	38.8	550 15	US-10-267-756-5
17	1230	38.8	559 16	US-10-858-271-28
18	1186.5	37.4	554 9	US-09-895-860-4
19	1186.5	37.4	554 15	US-10-377-072-4
20	1186.5	37.4	554 16	US-10-377-072-4
21	1179.5	37.2	571 10	US-09-931-836-23
22	1179.5	37.2	571 11	US-09-833-245-1090
23	1179.5	37.2	571 13	US-10-036-342-23
24	1179.5	37.2	571 13	US-10-036-041-23
25	1179.5	37.2	571 14	US-10-028-072-542
26	1179.5	37.2	571 14	US-10-035-855-23
27	1179.5	37.2	571 14	US-10-140-808-542
28	1179.5	37.2	571 14	US-10-121-049-542
29	1179.5	37.2	571 14	US-10-123-904-542
30	1179.5	37.2	571 14	US-10-140-470-542
31	1179.5	37.2	571 14	US-10-175-746-542
32	1179.5	37.2	571 14	US-10-176-918-542
33	1179.5	37.2	571 14	US-10-176-921-542
34	1179.5	37.2	571 14	US-10-227-884-210
35	1179.5	37.2	571 14	US-10-036-214-23
36	1179.5	37.2	571 14	US-10-137-865-542
37	1179.5	37.2	571 14	US-10-140-474-542
38	1179.5	37.2	571 14	US-10-035-719-23
39	1179.5	37.2	571 14	US-10-142-431-542
40	1179.5	37.2	571 14	US-10-143-114-542
41	1179.5	37.2	571 14	US-10-230-163-210
42	1179.5	37.2	571 14	US-10-036-160-23
43	1179.5	37.2	571 14	US-10-338-210
44	1179.5	37.2	571 14	US-10-142-419-542
45	1179.5	37.2	571 14	US-10-218-631-210

ALIGNMENTS

RESULT 1
US-10-023-515-2
; Sequence 2, Application US/10023515
; Publication No. US2002018266A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy A. J.
; TITLE OF INVENTION: Silos-Santiago, Immaculada
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/023, 515
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256, 369
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/279, 508
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 581

TYPE: PR1
ORGANISM: Homo sapiens
US-10-023-515-2

Alignment Scores:

Pred. No.:	4,94e-267	Length:	581
Score:	3079.00	Matches:	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	97.13%	Indels:	0
	13	Gaps:	0

US-10-023-515-3 (1-1746) x US-10-023-515-2 (1-581)

```
OY 1 ATGCCACGGGAACTTACTGATCTGCTTGCACAAATGSGTCTTTTCTGATTCCTGACGCC 60
Db 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuLeuLeuGlnPro 20
OY 61 CTGTTGGACACACAGACAGTGGGAAAAAAGCTGGGCTTCTGCTGAAGGCGCACAGAGAAC 120
Db 21 LeuLeuGlnYHisArgGlnTrpGlySerThrGlyProSerAlaGlnGlyProGlnArgAsn 40
OY 121 ACCAGGCTGGGATGATTCAGGGCAAGCAAGTCACTGTGCTGGGAAAGCCCTGCTGCTG 180
Db 41 ThrArgLeuGlyTrpIleGlnGlyGlnValThrValLeuGlySerProValProVal 60
OY 181 AACGATTCCTGGAGTCCCTTTGCTGCTCCCGGCTGGGATCCCTGGGATTTTACGAAAC 240
Db 61 AsnValPheLeuGlyValProPheAlaAlaProProLeuGlySerLeuArgPheThrAsn 80
OY 241 CCGAGCCCTGCATCGCCCTGGGATTAATTGCAAGAACCCACTCCTAATTTGTGTC 300
Db 81 ProGlnProAlaSerProTrpAsnLeuArgGlnAlaThrSerTrpProAsnLeuCys 100
OY 301 CTCGGAACCTCAGAGTGGCTCTTAAATCAACACATCTCTCAAGTGATTAACCGAAG 360
Db 101 LeuGlnAsnSerGlnTrpLeuLeuAspGlnHisMetLeuIleValHisTrpProCys 120
OY 361 TTCGAGTGTCAAGAGACTGCTACCTGAAACATCTATGCGCTGGCCGACGCGATACA 420
Db 121 PheGlyValSerGlnAspCysLeuTrpLeuAsnIleTrpAlaProAlaHisAlaAspThr 140
OY 421 GGCCTCAAGCTCCCGCTCTTGATGTGTTCCAGAGAGTGCTTCAAGATGGCTCAGCC 480
Db 141 GlySerTrpLeuProValLeuValTrpPheProGlyAlaPheLeuThrGlySerAla 160
OY 481 TCCATCTTTGATGGGTCGCGCTGCTGCTATGAGAACGTGCTGGTGTGGTCTGCCAG 540
Db 161 SerIlePheAspGlySerAlaLeuAlaAlaTrpGlnAspValLeuValValGln 180
OY 541 TACCGGCTAGAGAAATTTGGTTCTTCACACATGGGATCAGATGCTCCGGGAACTGG 600
Db 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpAspGlnHisAlaProGlyAsnTrp 200
OY 601 GCTTTCAGAGACAGAGTGGCTGCTGCTGCTGAGGATCCAGAGAACATCGAGTTCTTCG 660
Db 201 AlaPheIleAspGlnValAlaAlaLeuSerTrpValGlnIleAsnIleGlnPheGly 220
OY 661 GGGGACCCCAAGTCTGTGACCATCTTTGGCGAGTCCGGGGAGCCATAAGTCTTCTAGT 720
Db 221 GlyAspProSerSerValThrIlePheGlyIleSerAlaGlyAlaIleSerValSerSer 240
OY 721 CTTATACGTCTCCCATGGCCAAAGCTTATTCACAAAGCATGATGAGAGTGGGGTG 780
Db 241 LeuIleLeuSerProMetAlaIleGlyLeuPheHisIleValAlaIleMetGlnSerGlyVal 260
OY 781 GCCATCATCCCTTACCTGAGGCGCCATGATTAAGAAGAGTGAAGACCTGCAAGTGTG 840
Db 261 AlaIleIleProTrpLeuGlnIleHisAspTrpGlnIleSerGlnAspLeuGlnValVal 280
OY 841 GCACATTTCTGTGGTAAACATGCGTCAGACTCTGAGGCGCTGCTGAGGTGCTGAGGACA 900
Db 281 AlaHisPheCysGlyAsnAsnAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
```

```
OY 901 AAACCTCCAGGAGCTGTGACCTTCAGGCCAGAAAAAAGTCTTTCACCTGAGTGTT 960
Db 301 LysProSerLysGlnLeuLeuThrLeuSerGlnIleTrpLysSerPheThrArgValVal 320
OY 961 GATGGTCTTTCTTCTTAATGAGCCTTAGATCTATTGTCTCAGAAAGCATTTAAACA 1020
Db 321 AspGlyAlaPhePheProAsnGlnIleProLeuAspLeuLeuSerGlnIleAlaPheIleVal 340
OY 1021 ATTCCTTCATCATCGAGTCAATTAACACAGAGTGGCTTCTGCTGCTCCTATGAAGAG 1080
Db 341 IleProSerIleIleGlyValAsnAsnHisGlnCysGlyPheLeuLeuProMetLysGln 360
OY 1081 GCTCCTGAGATCCCTGAGTGGCTCCAAAGTCCCTTGGCCCTCCATGATACAAGATC 1140
Db 361 AlaProGlnIleLeuSerGlySerAsnLysSerLeuAlaLeuHisIleGlnIleHis 380
OY 1141 CTGACATCCCGCCTCAGTATTGTGACCTTGTGGCTAAATGAATCTTCATGACAAAGAC 1200
Db 381 LeuHisIleProProGlnTrpLeuHisIleValAlaAsnGlnTrpPheHisAspLysHis 400
OY 1201 TCCCTGATCGAAATCCGAGACAGTCTTCTGACTTGTGGAGATGTGTTCTTGTGGTC 1260
Db 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuGlyAspValPhePheValVal 420
OY 1261 CCTGCACTGATCAACAGCTCGATATCAAGAGATGCTGGTGCACCTGTCTACTTCTATGAG 1320
Db 421 ProAlaLeuIleThrAlaArgTrpHisArgAspAlaGlyAlaProValTrpPheTrpIle 440
OY 1321 TTTGGGACCGGCTCAGTGTCTTTGAAGACAGAAAGCGGCTTTTGTCAAAGCCGACAC 1380
Db 441 PheArgHisArgProGlnCysPheGlnAspThrLysProAlaPheValIleAlaPheHis 460
OY 1381 GCTGATGAAGTCCGCTTGTGTTGCTGGTGTGCTCTTCCGAAAGGGGAGATGTTATGTT 1440
Db 461 AlaAspGlnValArgPheValPheGlyGlyAlaPheLeuLysGlyAspIleValMetPhe 480
OY 1441 GAAGGAGCCAGGAGAGAGAAAGTTACTGAGCCGGAAGATGAATTAATCTGAGCTACC 1500
Db 481 GlnGlyAlaThrGlnGlnGlnIleLeuLeuSerArgLysMetMetLysTrpAlaThr 500
OY 1501 TTTGCTCAACCGGAGATCTTAATGGAACGACTGTCTGTGGCCACTTAATATGTG 1560
Db 501 PheAlaArgThrGlyAsnProAsnGlyAsnAspLeuSerLeuTrpProAlaTrpAsnLeu 520
OY 1561 ACTGAGCATACCTCCAGCTGGAATGGAACATGAGCCTCGGACAGAGACTGAAGAACG 1620
Db 521 ThrGlnGlnTrpLeuGlnLeuAspLeuAsnMetSerLeuGlyGlnArgLeuLysGlnPro 540
OY 1621 CGGATGATTTTGGACACGACCATCCCTGATCTGTGCTGCTGCTCCGACATGCTCCAC 1680
Db 541 ArgValAspPheTrpThrSerThrIleProLeuIleLeuSerAlaSerAspMetLeuHis 560
OY 1681 AGTCTCTTTCTTCTTAACTTTTCTCTCTCTCTCTCCAGCCTTTCTTTTGTGTGCT 1740
Db 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCysAla 580
OY 1741 CCT 1743
Db 581 Pro 581
```

RESULT 2
US-10-674-636-2
; Sequence 2, Application US/10674636
; Publication No. US20040086922A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A. J.
; TITLE OF INVENTION: 53010, A NOVEL HUMAN CARBOXYL ESTERASE
; TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
; FILE REFERENCE: 10448-122001
; CURRENT APPLICATION NUMBER: US/10/674,636
; CURRENT FILING DATE: 2003-09-29

; PRIOR APPLICATION NUMBER: US/10/023,515
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: 60/256,369
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: 60/279,508
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 581
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-674-636-2

Alignment Scores:

Pred. No.:	4,94e-267	Length:	581
Score:	3079.00	Matches:	581
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	97.13%	Indels:	0
DB:	15	Gaps:	0

US-10-023-515-3 (1-1746) x US-10-674-636-2 (1-581)

```

QY 1 ATGCCACAGGAGCTTACTTCATCTGCTTCAACAATGGTGGCTTTTCTTGATTCACAGCC 60
DB 1 MetProGlnGlyLeuTherSerSerAlaSerGlnTrpCyPhePheLeuIleLeuGlnPro 20
QY 61 CTGTGGGACACACAGACAGTGGGGAAAACTGGGCTTCTGCTGAAGAGCCACAGAGAAC 120
DB 21 LeuLeuGlnYHleAlaArgGlnTrpGlyLeuThrGlyProSerAlaGlnGlyProGlnArgan 40
QY 121 ACCAGGCTGGGATGGATTCAGAGGCAAGCAAGTCACTGTGCTGGGAAGCCCTGCTGCTG 180
DB 41 ThrArgLeuGlyTrpIleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
QY 181 AAGTGTCTCTGAGAGTCCCTTTGTGCTCCCGCTGGAGATCCCTGGCATTTACGAAC 240
DB 61 AenValPheLeuGlyValProPheAlaAlaProPheLeuGlySerLeuArgPheThrAn 80
QY 241 CCGAGAGCTGCATCGCCCTGGGATTAATTGGAGAAAGCACTCTTAATTTTGTGC 300
DB 81 ProGlnProAlaSerProTrpPhePheLeuArgGlnAlaThrSerTrpProLeuLeuGly 100
QY 301 CTCGCAAGACTCAGAGTGGCTGCTTAAATCAACACATGCTCAAGGTGATTAACCGGAA 360
DB 101 LeuGlnAenSerGlnTrpLeuLeuAenPheGlnHleMetLeuLysValHleTrpProLys 120
QY 361 TTGGAGTGTCAAGAGACTGCTTACCTGAACATCTATGCGCTGCGCAGCGCGATACA 420
DB 121 PheGlyValSerGlnAenPheLysLeuTrpLeuAenHleTrpAlaProAlaHleAlaPheTr 140
QY 421 GGCTCCAGAGCTCCCGCTTGTGATGTGTTCCAGAGAGTGGCTTCAAGACTGCTCAGCC 480
DB 141 GlySerLysLeuProValLeuValTrpPheProGlyGlyAlaPheLysThrGlySerAla 160
QY 481 TGCATCTTGAATGGGTCCCGCTGCTGCTTAAAGAGAGTGTGTTGTGCTGCTCAG 540
DB 161 SerIlePheAenPheGlySerAlaLeuAlaAlaTrpGlnAenPheValLeuValValGln 180
QY 541 TACCGGATAGAAATTTGTTTCTTCAACATGGAGATCAGATGCTCCGGGAGAACTGG 600
DB 181 TyrArgLeuGlyIlePheGlyPhePheTrpThrTrpPheGlnHleAlaProGlyAenTrp 200
QY 601 GCCTTCAAGAGACAGGTGCTGCTGTGCTGAGGCTCAGAGAAACATCGAGTTCTTCGT 660
DB 201 AlaPheLysAenPheGlnValAlaAlaLeuSerTrpValGlnLysAenHleGlnPhePheGly 220
QY 661 GGGAGACCCAGCTCTGTGACATCTTTGGCAGAGTCCCGGAGACCTAAGTGTCTTAAGT 720
DB 221 GlysAenProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
QY 721 CTTAATCTGTCTCCATGCGCAAGGCTTAATTCACAAAGCATCATGAGAGTGGGTG 780
  
```

```

DB 241 LeuIleLeuSerProMetAlaLysGlyLeuPheHleLysAlaIleMetGlnSerGlyVal 260
QY 781 GCCATCATCCCTTACTCGAGAGCCCATGATTTATGAGAAAGAGAGAGACTGCAAGGTGTT 840
DB 261 AlaIleIleProTyLeuGlnAlaHleAenPheTrpGlnLysSerGlnAenPheGlnVal 280
QY 841 GCACATTTCTGTGTAAACAATGCGTCAAGTCAAGTGGAGGCTGAGTGGCTGAGAGACA 900
DB 281 AlaHlePheCyGlyAenAenAlaSerAenPheSerGlnAlaLeuLeuArgCyLysAenGlyTr 300
QY 901 AAACCTCCAGAGAGCTGTGACCTCAGCCAGCAAGAAACAAAGCTTTTCACTGAGTGGTT 960
DB 301 LysProSerLysGlyLeuLeuThrLeuSerGlnLysThrLysSerPheThrArgVal 320
QY 961 GATGTGCTTTCTTTCTTAATGAGGCTCAAGATCTATGCTCAAGAAACATTTAAAGCA 1020
DB 321 AenGlyAlaPhePheProAenGlnProLeuAenPheLeuSerGlnLysAlaPheLysAla 340
QY 1021 ATTCCCTTCATGATCGAGTCAATTAACAGAGTGTGCTTCCGCTGCTGCTGCTGAGAG 1080
DB 341 IleProSerIleIleGlyValAenAenHleGlnCyGlyPheLeuLeuProMetLysGln 360
QY 1081 GCTCTGAGATCTCAGTGGCTCAACAAAGTCCCTTGCCTCATGTGATTAACAAACATC 1140
DB 361 AlaProGlnIleLeuSerGlySerAenLysSerLeuAlaLeuHleLeuIleGlnAenIle 380
QY 1141 CTCGACATCCCGCTCAGATTTTGAACCTTTGGGCTTAAGATTAATCTTCCATGACAGAC 1200
DB 381 LeuHleIleProGlnTrpIleuHleValAlaAenGlnLysPheHleAenPheLysVal 400
QY 1201 TCCCTGACTGAATCCGAGACAGTCTTGTGAGCTTGGATGATGTTGTTTGTGGTGC 1260
DB 401 SerLeuThrGlnIleArgAenPheLeuAenPheLysGlyAenPheValPhePheVal 420
QY 1261 CCTGACATGATCAACAGCTGATATCAACAGAGTGTGATGACCTGTGATTAATGAG 1320
DB 421 ProAlaLeuIleThrAlaArgTrpHleAenPheAlaGlyAlaProValTrpPheTrpGln 440
QY 1321 TTTCGACACCGGCTCAGTGGCTTTGAAGACAGAGACCGGCTTTTGTCAAGCCGACAC 1380
DB 441 PheArgHleAenPheGlnCyAenPheGlnAenPheTrpLysProAlaPheValLysAlaPhePhe 460
QY 1381 GCTGATGAATCGGCTTGTGTTGGTGTGCTTCCGAGAGGAGCAATGTTATGTTTC 1440
DB 461 AlaAenPheLysValArgPheValPheGlyAlaPheLeuLysGlyAenPheValMetPhe 480
QY 1441 GAAGAGCCACGAGAGAGAGAGATTACTGAGCCGAGAGATGATGAATATCTGGCTAAC 1500
DB 481 GlnGlyAlaThrGlnGlnGlnLysLeuLeuSerArgLysMetMetLysTrpTrpAlaThr 500
QY 1501 TTTCCTCAACCGGAGAAATCTTAATGGAACGACTGTCTGTGTGCTGCTTAATATCTG 1560
DB 501 PheAlaArgThrGlyAenProAenGlyAenAenPheLeuSerLeuTrpProAlaTyraAenLeu 520
QY 1561 ACTGAGAGTACTCAGCTGAGTGAATTAAGAGCTGCGGACAGAGACTCAAGAGACG 1620
DB 521 ThrGlnGlnTrpLeuGlnLeuAenPheAenMetSerLeuGlyGlnArgLeuLysGlnPro 540
QY 1621 CCGGTGATTTTGGACACAGACCAATCCCTGATCTGTGCTGCTGCTGCTGAGATGCTCAC 1680
DB 541 ArgValAenPheTrpPheSerThrIleProLeuIleLeuSerAlaSerAenPheLeuHle 560
QY 1681 AGTCTCTTTTCTTAACTTTCTGCTGTCTGCTGCTGAGCTTTTCTTTTGTGTTGCT 1740
DB 561 SerProLeuSerSerLeuThrPheLeuSerLeuLeuGlnProPhePhePheCyAla 580
QY 1741 CCT 1743
DB 581 Pro 581
  
```

RESULT 3
 US-10-757-262-46

```
; Sequence 46, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Eliasof, Scott D.
; APPLICANT: Siles-Santiago, Imaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2518, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MP103-00721R0NMNTM
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-757-262-46

Alignment Scores:
Pred. No.: 4,946-267 Length: 581
Score: 3079.00 Matches: 581
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.13% Indels: 0
Gaps: 0
DB: 16

US-10-023-515-3 (1-1746) x US-10-757-262-46 (1-581)
QY 1 ATGCCACAGGGACTTACTTCATCTGCTTCAACAATGAGTCTTTTCTGATTTCTCAAGCC 60
DB 1 MetProGlnGlyLeuThrSerSerAlaSerGlnTrpCysPhePheLeuLeuGlnPro 20
QY 61 CTGTTGGGACACAGACAGTGGGGAAGAACTGGGCGCTTGTGCTGAAGGGGACAGAGAAC 120
DB 21 LeuLeuGlnGlyAlaSerGlnTrpGlySerAlaGlnGlyProGlnArgAsn 40
QY 121 ACAGAGCTGAGATGATTCAGGGCAAGACACTGCTGCTGGGAAGCCCTGCTGCTGTG 180
DB 41 ThrArgLeuGlnGlyTrpLleGlnGlyLysGlnValThrValLeuGlySerProValProVal 60
QY 181 AACGTTCTCTGAGAGTCCCTTTGCTGCTCCCGGCTGGGATCCCTGGGATTACGAAC 240
DB 61 AsnValPheLeuGlyValAlaProPheAlaAlaProPheLeuGlySerLeuArgPheThrAsn 80
```

```
QY 241 CCGACGCTGACGCGCCCTGGGATTAATTGCGAAGACCACTCTACCTTAATTGTC 300
DB 81 ProGlnProAlaSerProTrpAsnLeuArgGlnAlaThrSerTrpProAsnLeuCys 100
QY 301 CTCAGAACTCAGAGTGCCTGCTTTAAGTCAACACATGCTCAAGCTGCATTACCCGAA 360
DB 101 LeuGlnAsnSerGlnTrpLeuLeuLeuAspGlnHisMetLeuValHisIleTrpLys 120
QY 361 TTCGAGAGTTCAGAGACTGCTCTTACCTGAACACTTATGCGCCGCCACGCGCATCA 420
DB 121 PheGlyValSerGlnAspCysLeuValLeuValIleValAlaProAlaHisIleAspThr 140
QY 421 GGCCTCAAGCTCCCGCTCTGCTGCTGCTTCCAGAGAGTGCCTTCAAGACTGCCTCAGCC 480
DB 141 GlySerTrpLeuProValLeuValTrpPheProGlyGlyAlaPheValThrGlySerHis 160
QY 481 TCCATCTTTGATGGTCCGCCCTGGCTGCTCATAGAGACGTGCTGTTGGTCTCCAG 540
DB 161 SerIlePheAspGlySerAlaLeuAlaIleTyrgIuAspValLeuValValValGln 180
QY 541 TACCGGCTAGGAATATTGTTGTTCTTCAACAATGGGATTCAGCATGCTCCGGGGAAC 600
DB 181 TyrArgLeuGlyIlePheGlyPhePheThrTrpTrpAspGlnHisAlaProGlyAsnTrp 200
QY 601 GCCTTCAAGACACAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 201 AlaPheValAspGlnValAlaAlaLeuSerTrpValGlnValAsnIleGluPheGly 220
QY 661 GGGGACCCCAAGCTCTGTACCAATCTTTGACGAGCTCCGGGAGCCATAGTCTTCTAGT 720
DB 221 GlyAspProSerSerValThrIlePheGlyGlnSerAlaGlyAlaIleSerValSerSer 240
QY 721 CTTATACGTCTCCCATGCGCAAGGCTTATTTCCAAAGCATCATGAGAGTGGGGTG 780
DB 241 LeuIleLeuSerProMetAlaValAspGlyLeuPheHisIleValIleMetGlnSerIleVal 260
QY 781 GCCATCATCTCTTACCTGAGGCGCCATGATTATGAAAGAGTGAAGACCTGACGCTGTT 840
DB 261 AlaIleIleIleProTrpLeuGlnAlaHisIleAspTrpGluValSerGlnAspLeuGlnVal 280
QY 841 GCACATTTCTGTGTAAACATATGCTGACACTCTGAGGCGCTGCTGAGTGCCTGAGACA 900
DB 281 AlaHisPheCysGlyValAsnAlaAlaSerAspSerGlnAlaLeuLeuArgCysLeuArgThr 300
QY 901 AAACCTTCAGAGAGCTGCTGACCTTCAGCCAGAAAGAAAGTCTTACCTGAGTGGTT 960
DB 301 LysProSerTrpGlyLeuLeuThrLeuSerGlnValSerTrpThrArgValVal 320
QY 961 GATGAGTCTTCTTCTTAATGAGGCTCTAGATTAATTGCTCAGAAAGCAATTTAAACA 1020
DB 321 AspGlyAlaPhePheProAsnGlnProLeuAspLeuLeuSerGlnValAlaPheValAla 340
QY 1021 ATTCTCTTCATCATGAGATCAATTAACACAGAGTGTGCTTCTGCTGCTTATGAAGAG 1080
DB 341 IleProSerIleIleGlyValAlaAsnAsnHisGlnCysGlyPheLeuLeuProMetCysGln 360
QY 1081 GCTCTGAGATTCCTCAGTGGCTCCAAAGATCCCTTGGCCCTCCATCTGATCAAAACATC 1140
DB 361 AlaProGlnIleLeuSerGlySerAsnLysSerIleuAlaLeuHisValIleGlnAsnIle 380
QY 1141 CTGCACATCCGCGCTCAGTATTTTGACCTTGTGCTAATGAATCTTCCATGACAGAC 1200
DB 381 LeuHisIleProProGlnTrpIleuHisValLeuValAlaAsnGlnTrpPheHisIleAspHis 400
QY 1201 TCCCTGACTGAATTCGAGACAGTCTTGTGACCTTGTGAGATGTGTTCTTTGTGCTC 1260
DB 401 SerLeuThrGlnIleArgAspSerLeuLeuAspLeuLeuGlyAlaPheValPhePheValVal 420
QY 1261 CCTGACATGATCAACGCTCGATATACAGAGATGCTGTGACACCTGTACTTCTATAG 1320
DB 421 ProAlaLeuIleThrAlaArgTrpHisIleArgAspAlaGlyAlaProValIleTrpPheTrpGlu 440
QY 1321 TTTGCGACCGGCGCTCAGTGTCTTTGAAGACAGAGCGGCTTTTGTCAAAGCCGAC 1380
```


QY 751 TTCCAAAGCCATCATGAGAGATGGGGTGGCCATCATCTTACCTGAGGCCCATGAT 810
 Db PhehlsyvalailemcccluserglyvalalailelleProtyldeuuliahisaer 331
 QY 811 TATGGAAGAGTGAAGACCTGCAAGGTGGTTCGACATTTTGTGGTAAACAATGGCTGAC 870
 Db TyrgluysserglulabpueglnvalalanhsphecygglYasnaenlaserar 351
 QY 871 TGTGAGGCCCTGAGGAGGCTGAGGAGCAAAACCTCCAAAGAGGTGTGACCCCTGAC 930
 Db SerglualaleuuarGyserlueuarglthlyrProserlysglueuueuthrueuser 371
 QY 931 CAGAAAACAAGCTTTTCACTGAGTGTGATGGTGTCTTTCTTAATGAGCCTTA 990
 Db GlnlystrlysserPhehthargvalValaaprglyalapherProhbngluproleu 391
 QY 991 GATCATTTGTCTCAGAAAGCATTTAAAGCAATTCCTTCATCATGAGAGTCAATTAACAC 1050
 Db AspleuleuserGlnlysalapheylsalalleProserllelleglyValaenbnhis 411
 QY 1051 GAGTGTGGCTTCGTGCTGCTATGAAGAGGCTCCGAGATGCTGAGTGGCTCCAAAG 1110
 Db GlnyGsglyPheleuuePrometlysglualaprogluvalleuserlYserenly 431
 QY 1111 TCCCTGGCCCTCATCTGATACAAAACATCTGACATCCGCTCAGATTTGACACTT 1170
 Db SerleuhalaleuhleuulleghnanielleuhlleleProProclnlyrleuuhlsleu 451
 QY 432 SerleuhalaleuhleuulleghnanielleuhlleleProProclnlyrleuuhlsleu 451
 QY 1171 GTGGTAAATGAATTAATCTTCATGACAGACACTCCCTGACTGAATCCGAGACGTTCTG 1230
 Db ValalaenGlnlyrPhehlsasplyshsSerleuthrGlnlleargasperleuueu 471
 QY 1231 GACTTGCTTGGAGATGTGTTCTTTGTGGTCCCTGACCTGATGACAGCTGATATACAGA 1290
 Db AspleuleuGlyAspvalPhePhevalValProalaLeuilethralarglyrthlsarg 491
 QY 1291 GATGCTGGGACACTGTACTTCTATGAGTTTGGGACCGGCTCAGAGCTTTGAGAGAC 1350
 Db AspalalaglyalaprovallyrPhehlyrgluphehghlsargProclnlyrshhegluabp 511
 QY 492 AspalalaglyalaprovallyrPhehlyrgluphehghlsargProclnlyrshhegluabp 511
 QY 1351 AGGAAGCCGGCTTTTGTCAAGCCGACACAGCTGATGAAAGTCGCTTTGTGGTGGTGT 1410
 Db ThrlyrProalaPheVallysalaphehlsalaleuabpGlnlyalArghevalPheglly 531
 QY 531 GCTTCTCTGAAGGGGACATTGTATGTTTGAAGGACACGAGAGAGAGAAAGTTACTG 1470
 Db AlaPheleuylgylAspillevalmetPheglnglyalathrGlnlylueuueu 551
 QY 551 AGCCGGAAGATGATAATTAATGAGGCTACTTTGTGCTGCAACGGGAAATCTAATGGGAGC 1530
 Db SerArglysmecmetlyertyrTrrpalaThrPhealaaGlyrthlyAsnProabnglyasn 571
 QY 1531 GACTGTCTCTGTGGCCAGCTTATATGACTGAGACGACTGACCTCCAGCTGAGCTTGAAC 1590
 Db AspleuleuPheleuThrProalaTrzhanleuThnglnlyrleuueuabpueuabn 591
 QY 592 MetSerleuGlyglahargleuylsgluproArgvalGlnpuehtrphrserThrllepro 611
 Db 1651 CTGATCTGTCTGCTCCGACATGCTGCAAGTCCCTTCTTCTTAATCTTCTCTCT 1710
 QY 612 LeuileleuserlaseraspeuhsseProleuserSerleuuthrPheuser 631
 QY 1711 CTCTTCAGACTTCTTTCTTTTGTGGTCTCT 1743
 Db 632 LeuleuGlnProPhePhePheCyalaapro 642
 RESULT 6
 US-10-451-168-92
 ; Sequence 92, Application US/10451168
 ; Publication No. US20040091969A1
 ; GENERAL INFORMATION:

/ APPLICANT: SMITHKLINE BEECHAM CORPORATION
 / APPLICANT: SMITHKLINE BEECHAM P.L.C.
 / APPLICANT: GLAXO GROUP LIMITED
 / TITLE OF INVENTION: NOVEL COMPOUNDS
 / FILE REFERENCE: GP50039
 / CURRENT APPLICATION NUMBER: US/10/451,168
 / PRIOR FILING DATE: 2003-11-12
 / PRIOR APPLICATION NUMBER: PCT/US01/49232
 / PRIOR FILING DATE: 2000-12-17
 / PRIOR APPLICATION NUMBER: 60/256,710
 / PRIOR FILING DATE: 2000-12-19
 / PRIOR APPLICATION NUMBER: 60/257,048
 / PRIOR FILING DATE: 2000-12-20
 / PRIOR APPLICATION NUMBER: 60/260,482
 / PRIOR FILING DATE: 2001-01-09
 / PRIOR APPLICATION NUMBER: 60/264,922
 / PRIOR FILING DATE: 2001-01-30
 / PRIOR APPLICATION NUMBER: 60/266,797
 / PRIOR FILING DATE: 2001-02-06
 / PRIOR APPLICATION NUMBER: 60/276,988
 / PRIOR FILING DATE: 2001-03-19
 / PRIOR APPLICATION NUMBER: 60/281,535
 / PRIOR FILING DATE: 2001-04-04
 / PRIOR APPLICATION NUMBER: 60/289,622
 / PRIOR FILING DATE: 2002-06-28
 / NUMBER OF SEQ ID NOS: 110
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 92
 / LENGTH: 581
 / TYPE: PR
 / ORGANISM: Homo sapiens
 US-10-451-168-92
 Alignment Scores:
 Pred. No.: 1,07e-224 Length: 581
 Score: 2607.50 Matches: 499
 Percent Similarity: 92.67% Conservative: 7
 Best Local Similarity: 91.39% Mismatches: 25
 Query Match: 82.26% Indels: 15
 DB: 15 Gaps: 3
 US-10-023-515-3 (1-1746) x US-10-451-168-92 (1-581)
 QY 34 TGGTGCTTTTCTGATGTTCTCAGGCCCTGTTGGGACACAGACAGTGGGAAACTGGG 93
 Db 14 Trrpalailetrpvalleuulalaapro-----Thrlysgly 25
 QY 94 CCTTGTGCTGAAGGGCCACAGAGAAACACAGCTGGATGATTCAGGCAAGCAAGTC 153
 Db 26 ProserAlaglnlyrProglmrganbthargleuGlyTrrpilleGlnlysglnval 45
 QY 154 ACTGTGTGGAGAGCCCTGTGCTGTGAACGTTGCTTCGAGAGTCCCTTTGTGCTGCC 213
 Db 46 ThrValleuGlySerProvalProvalaenValPheleuGlyvalProPhealalaapro 65
 QY 214 CCGCTGGGATCCCTCGATTTTGAACCCGAGCGCTGATGCGCTGGGATTAATTGGGA 273
 Db 66 ProleuGlySerleuargPhehthrsanProglnProalaSerProtrpabnleuarg 85
 QY 274 GAAGCACTCTTACCTTAATTGCTCCAGAACTCAGAGTGGCTGCTTAGATCAA 333
 Db 86 GlnlathrSerlyrProhbnleuGlyleuGlnanSerGlntrPheleuueuabpGln 105
 QY 334 CACATGCTCAAGTGCATTAACCGAAATTTGGAGTGCAGAAAGTGCCTTACTGAAC 393
 Db 106 HlsmeleuylvalahierlyrProlysphegllyValserGlnabpCyserlyrleuabn 125
 QY 394 ATCTATGGCCCTGGCCACGCGCATACAGGCTCCAGGCTCCCGCTTGGTGTGGTCCCA 453
 Db 126 lleyrAlaProalaHlsalaspthrglyserlyrleuProvalleuvaltrpPhepro 145
 QY 454 GGAGGCTTCAAGACTGCTCAGCTTCATCTTGAATGGTCCGCTGCTGCTCAT 513

[illegible]

```

QY      1573  CTCGACGTGGACTTGAACATAAGCTTGAGACAGAGCATCAAGAAACCGCGGATGATTTT 1632
Db      526   LeuGlnLeuAapLeuAasMetSerLeuGlyGAlnArgLeuYbGInProARgAspVal 545

QY      1633  TGCACGACGACCATCCCC 1650
Db      546   TrpValThrGlyTyPro 551

RESULT 7
US-10-094-749-2375
; Sequence 2375, Application US/10094749
; Publication No. US20030219741A1
GENERAL INFORMATION:
APPLICANT: ISOCHI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2375
LENGTH: 525
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2375

Alignment Scores:
Pred. No.:          1,4e-224           Length:       525
Score:             2606.00            Matches:      503
Percent Similarity: 88.77%            Conservative: 3
Best Local Similarity: 88.25%          Mismatch:     6
Query Match:        82.21%             Indels:       58
DB:                 15                Gaps:         2

US-10-023-515-3 (1-1746) x US-10-094-749-2375 (1-525)

QY      34    TGTGTCCTTTTCCGATTCCTCCAGCCCCTGTGGAGACAGACAGTGGGGAAAACTGGG 93
Db      14    TrpAlaIleTrpValLeuAlaAlaPro-----ThrLyseGly 25

QY      94    CCTCTGTGTAAGGCCACAGAGAACACCAAGCTGGAGATTCAGGGCAGCAAGTC 153
Db      26    ProSerAlaGluGlyProGlnArgAsnThrArgLeuLysTrpIleGlnGlybSGlnVal 145

QY      154   ACTGTGCTGGAGACCCCTGTGCTGCTGTGAAGCTGTTCTCGAGAATCCCTTTGCTGCTCC 213
Db      46    ThrValLeuGlySerProValProValAsnValPheLeuGlyValProPheAlaAlaPro 65

QY      214   CCGCTGGGATCCCTGCGATTTCAGAACCCGACGCTTGATCGCCCTGGATTACTGCGA 273
Db      66    ProLeuGlySerLeuArgPheThrAsnProGlnProAlaSerProTPRAspAsnLeuArg 85

```



```

; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 196
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-196

Alignment Scores:
Pred. No.:      8,47e-224      Length:      581
Score:          2597.50        Matches:      497
Percent Similarity: 95.07%      Conservative: 4
Best Local Similarity: 94.31%    Mismatches:  17
Query Match:    81.94%         Indels:       9
DB:             15             Gaps:         3

US-10-023-515-3 (1-1746) x US-10-114-270-196 (1-581)
QY 91 GGGCCCTTGTGCGAAGGGCCAGAGGAAACACAGGCTGGATTCAGGGCAAGCAA 150
DB 27 Glycerolaldehyde 1-propanolamine triethylamine 1-ylglycin 46
QY 151 GTCACTGTGCTGGAGAGCCCTGTGCTGTGAACGTTCCTCGAGTCCCTTTGTGCT 210
DB 47 Valthirvalineuglyserprovalphenalenvalphenuglyvalprohehala 66
QY 211 CCCCCGTGGAGTCCCTGCGATTTCAGAACCCGACGCTGCATCGCCCTGGATTAATTG 270
DB 67 ProProleuglyserleuarghetheranProglinProalaserProtrapsanleu 86
QY 271 CGAGAGGCACTCCCTACCTAATTGTGCTCCAGAACTCAGAGGCTGCTCTTAGAT 330
DB 87 ArgglualathserlytrprosanleucylseuglnanserGlntrpleuuleuamr 106
QY 331 CAACACATGCTCAAGGTGATTAACCGAATTCGAGGTGCAGAGACTGCTTACCTG 390
DB 107 GlnhismetleuValhithytrProlysbheglValserGluamrPcylseutyrlieu 126
QY 391 AACATCTATGCCCTGCCCAACGCCGATACAGGCTCCAGCTCCCGTCTTGTTGTTTC 450
DB 127 AsniletyralaProalathialaerThirgylserlyleuProvalleuValtrPhe 146
QY 451 CCAGAGAGGCTTGAAGCTGCTCAGCTCAGCTCATCTTTGATGGGCTGCCGCTGCTGC 510
DB 147 ProgllyalalaphelyserhithyseralaserllephesrPglySerlaleuulala 166
QY 511 TATGAGAGCTGCTGTTGTGCTGCTCAGTACCGGCTAGAGAAATTTGGTTCTTCAAC 570
DB 167 TyrgluamrvalleuValvalValglntyrzargleuglylalephelybhehethr 186
QY 571 ACATGGGATCAGCATGCTCCGGGGAACCTGGGCTTCAAGAACCAAGTGGCTGCTGTCC 630
DB 187 ThrtpramrGlnhialaProgllyantrpalaPhelelyasrPglValalaaleuSer 206
QY 631 TGGGTCAGAGAGAATGAGTTCTTCGGTGGGGAACCCAGCTCTGTGACCACTTTGGC 690
DB 207 TrpvalGlnlyAsnilleGlnbhehnegllygllyamrProserSerValthrllephely 226
QY 691 GAGTCGGGGGAGCCATAGTGTCTTACTTATAGTGTCTCCATGGCCAAAGGCTTA 750
```

```

DB 227 GluseralaglyalalaserValserleuileuSerPromerAlalygllyeu 246
QY 751 TTCCCAAGGCATATAGAGAGTGGGTCATCATCCCTTACCTGAGGCCATATAT 810
DB 247 PhehlslyballemeGluserlValalalleleProlyrleuglnlahlsamr 266
QY 811 TATGAGAGAGTGAAGCACTCAGAGTGGTGCACATTTCTGTGTAAATAGCTCAAC 870
DB 267 TyrglnlySerGluamrleuGlnValvalalansphecylglyasrshalsaser 286
QY 871 TCTGAGCCCTGCTGAGGTGCTGAGCAAAACCTTCAAGAGACTGACCTGACG 930
DB 287 SerGlnlaleuileuArgCyleuamrThryrProserlyGlnleuThreuser 306
QY 931 CAGAAAACAAAGTCTTTACTTCGAGTGTGATGCTTTCTTCTTAATAGAGCTTA 990
DB 307 GlnlyethrlybSerPheThrArgValalaspGlyalalapherProasngluProleu 326
QY 991 GATCTATTGTCTCAGAAAGCAATTTAAAGCAATTCCTTCATCATCGAGTCAATACAC 1050
DB 327 AspleuileuSerGlnlyalalaphelyalalleProserlleileglyValashanshls 346
QY 1051 GAGTGTGCTTCTGTGCTGCTATGAAGAG-----GCTCCTGAGATCTGAGTCTCC 1104
DB 347 GlucylglyPheleuLeuPromerValArglleuulalaValhithrAlathrProser 366
QY 1105 AACAAAGTCCCTGCTCC-----CATCGATACAAACATCCTGACATCT 1149
DB 367 AsnArgpAlalalaleuallaserThralaglyhlsPhehlsargArg-----Hstlle 384
QY 1150 CCGCTCAGATTATTTGACCTTGAGCTTAATGATCTTCATGACCAAGCACTCCCTGACT 1209
DB 385 ProProglntyrleuHlsleuValalalansglutyrrhenlsamrlyhlsierleuthr 404
QY 1210 GAAATCCGAGACAGTCTTTGAGTCTGCTGAGATGTGTTCTTTGTGCTCCCTGACTG 1269
DB 405 GlulleargamrSerleuLeuamrleuGlnlyamrValPhePheValvalProalaleu 424
QY 1270 ATCAGAGTCGATATCAAGAGATGCTGTCACCTGTCTACTTATAGATTTCGGAC 1329
DB 425 lIethrAlaarglyrYhlsArgamrAlaglyalalProvaltyrPheTyrglnbhehthrs 444
QY 1330 CCGCTCAGTCTTGAAGACAGAACCGGCTTTTGTCAAGCGGACCAAGCTGATGA 1389
DB 445 ArgProglntyrPheGluamrThryrProalalPheVallybAlaamrshlsalsrPgl 464
QY 1390 GTCCGCTTGTGTGCTGCTGCTCTTCTGAAGGGGACATTTGTTATGTTGAAAGGACC 1449
DB 465 ValArgPheValPhegllyalalPheleuylglyamrilleValmecPheglndlyla 484
QY 1450 ACGAGAGAGAGAGATTTCAGAGCCGAGAGATGAATCTGGGCTACCTTGTCTGCA 1509
DB 485 ThrglnlyGlnlyleuLeuSertrglysmeclelytyrTrpAlathrPhehlsArg 504
QY 1510 ACCGGGAATCTTAATGGAGAGCACTGTCTGTGGCCAGCTTAATCTGATGAGCAG 1569
DB 505 ThrglyamrProamnglyamrshamrleuSerleutrpAlatyrrAsnleuThGlnGln 524
QY 1570 TACCTTCAGCTGGAATTGAACATGAGCTTGAGACAGAGACTCAAGAACCCGGGTGAT 1629
DB 525 TyrgleuglnleuAmrleuamrmeSerleuGlnArgleuylGlnProArgAmr 544
QY 1630 TTTTGGACCAAGACCATCC 1650
DB 545 ValtrpValthnrglytyrPro 551

RESULT 9
US-10-381-898-2
; Sequence 2, Application US/10381898
; Publication No. US2004008687A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valda; BAUGHN, Mariah R.;
```

APPLICANT: BOROMSKI, Mark L.; DING, Li;
APPLICANT: DUGGAN, Brendan; ELIOT, Vicki S.;
APPLICANT: GANDHI, Ameena R.; GRIFFIN, Jennifer A.;
APPLICANT: HAPPA, April J.A.; ISON, Craig H.;
APPLICANT: KHAN, Farrah A.; LAU, Preeti G.;
APPLICANT: LEE, Ernestine A.; LU, Dying Aina M.;
APPLICANT: ARVIZU, Chandra S.; POLICKY, Jennifer L.;
APPLICANT: RAMKUMAR, Jayalaxmi; RING, Huijun Z.;
APPLICANT: SANJANALA, Madhusudan M.;
APPLICANT: TANG, Y. Tom; TRIBOULEY, Catherine M.;
APPLICANT: CHAWLA, Nandinder K.; WALSH, Roderick T.;
APPLICANT: WARREN, Bridget; XU, Yuming;
APPLICANT: YANG, Junming; YAO, Monique; YUE, Henry
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-023 USN
CURRENT APPLICATION NUMBER: US/10/381, 898
CURRENT FILING DATE: 2003-10-17
PRIOR APPLICATION NUMBER: PCT/US01/30662
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/236,947
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: US 60/238,864
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/242,323
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/247,581
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/249,519
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/252,834
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/250,567
PRIOR FILING DATE: 2000-11-30
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 618
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7473645CD1
US-10-381-898-2

Alignment Scores:
Pred. No.: 5,596-223 Length: 618
Score: 2588.50 Matches: 491
Percent Similarity: 94.62% Conservative: 1
Best Local Similarity: 94.42% Mismatches: 5
Query Match: 81.66% Indels: 23
DB: 15 Gaps: 1

US-10-023-515-3 (1-1746) x US-10-381-898-2 (1-618)
QY 91 GGGCTTGTGCGAGGGGCAAGAGAACACAGGCTGGATGATTGAGGGCAAGCA 150
DB 92 GtlyProserAlagInglyProGlnArghaenThrArgleuGlyTyrIleGlnIlyyBgln 111
QY 151 GTCACTGTGCTGGAGAGCCCTGTGCTGGAAGCTGTTCTCGAGTCCCTTTGCTGCT 210
DB 112 ValThrValIleuGlySerProValProValaIenValPheIleuGlyValProPheAlaIa 131
QY 211 CCCCCGTGGGATCCCTGGATTTTACGAACCCGAGCTGCATCGCCCTGGAGTAACCTTG 270
DB 132 ProPheIleuGlySerIleuArghPheThrArpProGlnProAlaSerProThrPheIleu 151
QY 271 CGAAGAGCACTCTCTACCTTAATTTGGCTCCCGAATCTCAGAGTGGCTGCTTAAAT 330
DB 152 ArgGlnAlaThrSerIlyrProaIenLeuCySleuGlnaIenSerGlyIlyrIleuLeuIen 171
QY 331 CAACCATGCTCAAGTGATTAACCGAATTCGAGAGTGTCAAGAGACTGCTTAACTG 390
DB 172 GlnHISMetIleuIlyValaIleIlyrProIlySphGlyValSerGlyIlyrPheIlyrIleu 191

QY 391 AACATCTATGCGCTGCCCCACGCCGATACAGGCTCCAAAGTCCCGTCTGGTGGTTC 450
DB 192 AenIleTyrAlaProAlaIleIleAlaSerThrGlySerIlyrSleuProValIleuValIlyrPhe 211
QY 451 CCAAGAGGTGCTTCAAGACTGGCTGCACGCTTCATCTTTGATGGTCCGCTGCTGCC 510
DB 212 ProGlyGlyAlaPheIlyrThrGlySerAlaSerIlePheAepGlySerAlaIeuaIaIa 231
QY 511 TATGAGACGTGCTGGTGGTGGTCTCCAGATACCGGCTAGAAATATTTGGTCTTCAAC 570
DB 232 TyrGluAepValIleuValaIleValaGlnIlyrArgleuGlyIlePheGlyPhePheThr 251
QY 571 ACATGGGATCAGCATGCTCCGCGGAACCTGGGCTTCAAGAACCAAGTGGCTGCTGCC 630
DB 252 ThrTrpAepGlnHISAlaIleProGlyAenThrAlaPheIlyrAepGlnValaIleAlaIeuser 271
QY 631 TGGGTCCAGAAAGACATCGAGTCTTCGATGGGGAACCCAGCTGTGACATCTTTGGC 690
DB 272 TrpValGlnIlyrAenIleGlyIlyrPheGlyIlyrAepProSerValThrIlePheGly 291
QY 691 GAGTCCCGGGAGCCATAAGTGTCTTAATCTGTTCCCATGGCCAAAGCTTA 750
DB 292 GluSerAlaGlyAlaIleSerValSerSerIleuIleuSerProMetAlaIySglyIleu 311
QY 751 TTCACAAAGCATGATGAGAGTGGGGTGGCCATCATCCCTTACCTGGAGGCCATGAT 810
DB 312 PheHISlyrAlaIleMetGlySerGlyValaIleIleProIlyrIleuGlnAlaHISaP 331
QY 811 TATGAGAGAGTGAAGACCTGAGGTGGTGGACATTTCTGTGTAAACATCGTCAGAC 870
DB 332 TyrGlnIlyrSerGlyAepIleuGlnValaIleAlaHISphecYsglyAaIenAlaIeSerAep 351
QY 871 TCTGAGGCCCTGCTGAGGTGCTTGAAGACAAACCTCCAGAGAGCTGTGACCTCAGC 930
DB 352 SerGlnAlaIleuIleuArgCylSerIlyrThrIlyrProSerIlyrGlyIleuIleuThrIleuSer 371
QY 931 CAGAAAGCAAGTCTTTCATCTGAGTGGTGAATGGTCTTCTTCCTAATAGGCTCA 990
DB 372 GlnIlyrThrIlyrSerPheThrArgValaIlePheGlyAlaPhePheProaIenIlyrProIleu 391
QY 991 GATCTATTGTCTCAGAAAGCATTTTAAAGCAATTCCTTCATCATGCGAGTCAATACAC 1050
DB 392 AepIleuIleuSerGlnIlyrAlaPheIlyrAlaIleProSerIleIleIleGlyValaIenHIS 411
QY 1051 GAGTGTGGCTTCTGCTGCTGCTATGAGAGAGGCTCTGAGATCTGATGGCTTCAAAG 1110
DB 412 GluCySglyPheIleuIleuProMet----- 419
QY 1111 TCCCTTGCCCTTCATCTGATACAAACATCTCGACATTCGCCCTCAGATTTGCACTT 1170
DB 420 -----HisIleProProGlnIlyrIleuHISleu 428
QY 1171 GNGCTAAATGAATTACTTCATGACAGACACTCCGAGCTGAATTCGAGACAGTCTTGTG 1230
DB 429 ValAlaAenGlyIlyrPheHISaepIlyrHISerIleuThrGlnIleArgAepSerIleuIleu 448
QY 1231 GACTTGCTTGAAGATGTGTTCTTTGTGGTCCCTGCACTGATCAAGCTGATATACAGA 1290
DB 449 AepIleuIleuGlyAepValPhePheValaIleProAlaIleuIleThrAlaArgIlyrHISArg 468
QY 1291 GATGCTGGTGCACCTGTCTACTTCTATAGATTCCGACACCGGCTCAGAGCTTTGAAGC 1350
DB 469 AepAlaGlyAlaProValIlyrPheIlyrGlyIlyrPheArgHISArgProGlnIlyrPheGlyIlyr 488
QY 1351 ACGAAGCGGCTTTGTGTAAGCGGACGACGCTGATGAAGTCCGCTTGTGTGGTGGT 1410
DB 489 ThrIlyrProAlaPheValaIlyrAlaIlePheHISAlaAepGlyValaIlyrPheValaIleGlyIly 508
QY 1411 GCTTCTCGAAGGAGGACATTTATATGTTGGAAGAGCCAGAGAGAGAAAGTTACTG 1470
DB 509 AlaPheIleuIlyrGlyAepIleIleValaIlePheGlyIlyrAlaThrGlnGlyIlyrIlyrIleu 528

QY 1471 AGCGGAAGATGATGAATACTAGGCTACTTTGGCTCGAACCGGGAATCTTAATGGGAAC 1530
DB 529 Seratglysmetmetelytrtrpalaatrphelaabgthrghlyasnproahsnilyasn 548
QY 1531 GACCTGTCTTGTGGCCAGCTTATAATCTGACTAGACGATCACTTCAGCTGGACTTGAAC 1590
DB 549 Aspleuserleutrpolaatyranleuthrghlgnlyrleuglnleuasnpleuasn 568
QY 1591 ATGAGCCTCGGACAGAGACTCAAAAGACCGGGGTGAGATTTTGGACCGACCACTCCCC 1650
DB 569 Metserleuaglyglnhrghleuylseqluproarigaabpvaltrpvalthrghlytrpro 588

RESULT 10
US-10-104-047-2219
; Sequence 2219, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2219
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2219

Alignment Scores:
Pred. No.: 5,936-211 Length: 469
Score: 2454.00 Matches: 467
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.57% Mismatches: 0
Query Match: 77.41% Indels: 0
DB: 15 Gaps: 0

US-10-023-515-3 (1-1746) x US-10-104-047-2219 (1-469)
QY 337 ATGCTCAAGTGCATTAATCCGGAATTCGAGTGTGAGAAAGACTGCTCTGAACATC 396
DB 1 Metleuylsvalhistrproulyserlyphelgylaserghlyuapcybleuthrleuasnile 20
QY 397 TATGGCCTGCCCGCCAGCGGATACAGAGCTCCCGCTTGTGTGTGTCCAGGA 456
DB 21 Tyrhlaproahlaahlaabpthrghlyserlybleuaprovalleuvaltrpbeprgily 40
QY 457 GGTGCTTCAGAGACTGGCTCAGCTTCATCTTTGATGGGTCCGCCCTGGCTGCCATGAG 516
DB 41 Glylaapheulythrghlyserlyaserillepheabpghlyserlaileuvalaalyrghu 60
QY 517 GACGCTGT 576
DB 61 Aspvallleuvalvalvalvalglnlyrghleuyllephelgylphepethrthrtrp 80
QY 577 GATCAGATGCTCCGCGGGAATCGGCTTCAGAGACCAAGTGGCTGTCTGTCTGGGTC 636
DB 81 Aepghnhtbalaproglyasantrpalaapheulyabpghlnvalaialaleuasertrpval 100
QY 637 CAGAAAGAACATCGAGTTCTTGGTGGGAGACCCAGCTCTGTGACCATCTTTGGCGAGTCC 696
DB 101 Glnlysaaniileglnpbeprghlyglaabprousersevalthrillephelgylgubser 120
QY 697 GCGGAGGCGATTAAGTTTCTTAAGTCTTAAGTCTTAAGTCTTAAGTCTTAAGTCTTAAGT 756
DB 121 Alaelyalalieserlyaserleuyleuaserprometalaubelyleuhenhs 140
QY 757 AAAGCATCATGAGAGTGGGTGGCCATCATCTCTTACTGGAGGCCCATGATTATGAG 816
DB 141 Lysalalilemetgluserghlyvalalalietleleprotyrleuglnahlaahpbyrghu 160

QY 817 AAGATGAGACTGAGAGTGGTGGTGCATTTCTGTGTAACTAATGCTCAGACTGTAG 876
DB 161 Lysserghlyuapbeprghlnvalaialahisrhecsghlyasnabnalaaserasprerghu 180
QY 877 GCCCTGCTAGAGTGTCTGAGGACAAAGACCTTCAGAGAGTGTGACCTTCAGCCAGAA 936
DB 181 Alaileuwaahcghlyeuhargthrlyserproserlysgileuethrleuaserghlnly 200
QY 937 ACAAGTCTTCACTCGAGTGGTGTGATGTGTCTTCTTCTTAATGAGCCTTGATCTA 996
DB 201 Thrlyserpethrtrpvalvalaspghlyalabepheproahnglnprouleuasnpleu 220
QY 997 TTGTCTCAGAAAGACTTTAAAGCAATTCCTTCATCATCGAGTCAATTAACACAGAGTGT 1056
DB 221 Leuserghlnlybalaapheulyalaleproserilleelgylvalasnserhileghlycys 240
QY 1057 GCGTCTCTGTGCTGCTATGAGAGGCTCTGAGATCTCAAGTGGCTCCAGAGTCCCTT 1116
DB 241 Glypheleuaproumetlyeglnabpghlnabpghlnleuaserghlyserasnlyserleu 260
QY 1117 GCCCTCCATCTGATTAACAAACATCTGACATCCCGCTCAGTATTGGACACTTGTGGCT 1176
DB 261 Alaileuhtleuileglnasnilleuhtalileproproghlnlyrleuhtleuvalala 280
QY 1177 AATGAATCTTCATGACAGACACTCCCTGACTGAATCCGAGACAGTCTTGGACTTG 1236
DB 281 Aenglnlyrphelhaabpbyhahiserleuthrghlnleahgabserleuasnpleu 300
QY 1237 CTGGAGATGTGTCTTGT 1296
DB 301 Leughlyabpvalpbeprghlnvalprohlaileuthrilaargyghlaabpvala 320
QY 1297 GGTGACCTGTCTACTTTATGAGTTTGGGACCGGCTCAGTGTGTGAAGACAGAG 1356
DB 321 Glylaaprovaltyrphelyrghlypbeprghlnabpghlnabpghlnabpghlnabpghln 340
QY 1357 CCGGCTTTTTCAGAGCCGACAGCTGATGAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGT 1416
DB 341 Prohlaaphevallybalaasphilaahlaabpghlnvalabpbeprvalpbeprghlyalabhe 360
QY 1417 CTGAAGGGGACATTTGATTTGATGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1476
DB 361 Leuylsglyabpillevalmetpheghlncllyalatrghlgnlyrghlyleuaserarg 380
QY 1477 AAGATGATGAATACTGGGCTTACCTTTGTCTGGAACCGGGAATCTTAAGGGAAGACTTG 1536
DB 381 Lysmetcelytyrtrpalaatrphelaargthrghlyasnproahnglyasnabprouleu 400
QY 1537 TCTGTGTGGCAGCTTAATCTGACTGAGGAGTACTCCAGCTGAGACTTGAACATGAGC 1596
DB 401 Serleuthrproahlaatyranleuthrghlgnlyrleuglnleuasnpleuasmetseser 420
QY 1597 CTGGAACAGACTCAAGAAACCGGGGTGAGATTTTGGACCAAGCCATCCCTGATC 1656
DB 421 Leughlyghlnabpghlnabpghlnabpghlnabpghlnabpghlnabpghlnabpghln 440
QY 1657 CTGTGTGCTCTCGACATCTCCAGAGTCTCTTCTTCTTAACTTCTCTCTCTCTCTCTC 1716
DB 441 Leuserlaaserasmetleuhtleuaserproleuaserleuthrilleuaserleu 460
QY 1717 GAGCCTTCTTTTCTTTTGTGTCTCT 1743
DB 461 GlnprouphneprhepneCysalaapro 469

RESULT 11
US-10-233-933A-2
; Sequence 2, Application US/10233933A
; Publication No. US20040214177A1
; GENERAL INFORMATION:
; APPLICANT: Yamashita, Tetsuro
; APPLICANT: Miyazaki, Masao
; TITLE OF INVENTION: CAT KIDNEY DISEASE MARKER


```
FILE REFERENCE: SHIG FP02US006
CURRENT APPLICATION NUMBER: US/10/233, 933A
CURRENT FILING DATE: 2002-09-03
PRIOR APPLICATION NUMBER: JP2002-057908
PRIOR FILING DATE: 2002-04-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 542
TYPE: PRF
ORGANISM: Felis catus
US-10-233-933A-2

Alignment Scores:
Pred. No.: 8, 17e-169 Length: 542
Score: 1985.00 Matches: 373
Percent Similarity: 83.27% Conservative: 60
Best Local Similarity: 71.73% Mismatches: 85
Query Match: 62.62% Indels: 2
DB: 16 Gaps: 2

US-10-023-515-3 (1-1746) x US-10-233-933A-2 (1-542)
QY 91 GGGCCCTTCTGTAAGGCGCAGAGAAACACGAGCTGGAGTGAATTCAGGCGAAGCA 150
DB 22 GTPProAlaAlaApriAlaProValArgSerThrArgLeuGlyTrpValArgGlyGln 41
QY 151 GTCACTGTCTGGGAAGCCCTGTCTGTGAACGTCTTCTCGAGTCCCTTGTCTGCT 210
DB 42 ThrThrValLeuGlySerThrValProValAsnMetPheLeuGlyLeuProGlyAlaAla 61
QY 211 CCCCGCTGGGATCCCTGGATTATAGAACCCGAGCTGCATCGCCCGGATTAAGTCTG 270
DB 62 ProProLeuGlyProLeuArgPheLeuGlnProGlyProAlaLeuProGlyAsnMetPhe 81
QY 271 CGAAGACCACTCTTACCTTAATTTGTGCTCCAGAACTCAGAGTGGCTGCTTAAGT 330
DB 82 ArgAsnAlaThrSerLeuProGlySerLeuGlnPheGlnApriLeuGlyTrpValArg 101
QY 331 CAACACATCTCAGAGTGAATTAACCGAAATTCGAGTGTCAAGAGATGCTCTTAAGT 390
DB 102 GlnHsValLeuLeuValArgGlyProGlyLeuGlnAlaSerGlnAspGlyLeuGlyLeu 121
QY 391 AACATCTATGCGCTGCCAGCGGATAGAGCTCCAGACTCCCGCTGTGTGGTTC 450
DB 122 AsnIleTyrAlaProAlaHsAlaAspAsnGlySerAsnLeuProValMetValTrpPhe 141
QY 451 CGAAGAGTCTTCAAGACTGAGCTCCATCTTGAATGAGTCCGCGCTGAGCTGCC 510
DB 142 ProGlyGlyAlaPheLeuMetGlySerAlaSerSerPheAspGlySerAlaLeuAlaAla 161
QY 511 TATGAGAGCTGCTGTGTGTCTGCTCAATACCGGCTAGAAATATTGGTTCCTTACC 570
DB 162 TylGlnAspValLeuLeuValThrThrGlnTyrArgLeuGlyIlePheGlyPhePheAsp 181
QY 571 AGATGGATCAGACGCTCCGCGGAACTGGGCTTCAAGAGCAAGTGGCTGTCTGTCC 630
DB 182 ThrGlyAspGlnHsIleAlaArgIleAsnTrpAlaLeuLeuAspGlnAlaAlaAlaLeuThr 201
QY 631 TGGGTCAGAGAAACATCGAGTTCTTCGCTGGGGAACCCAGACTCTGTGACATCTTTGAC 690
DB 202 TrpValArgAspAsnIleGlnPhePheGlyGlyAspProArgSerValThrIlePheGly 221
QY 691 GAGTCGCGGGAAGCCATAAGTCTTACTTATATCTGTCTTCCATGCCCAAGGCTTA 750
DB 222 GlnSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaGlnGlyLeu 241
QY 751 TTCCCAAGGCAATCATGAGAGTGGGGGCGCATATCCCTTAAGTCTG---GAGGCCAT 807
DB 242 PheHsIleValAlaIleMetGlnSerGlyValAlaIleLeuProLeuLeuMetArgProPro 261
QY 808 GATTATGAGAAAGTGAAGCTGAGTGTGTGACATTTCTGTGTGAACATGCGCTGA 867
DB 867
```

```
DB 262 GlnAspGlnArgIleValSerPheLeuGlnValLeuAlaArgIleCysGlyCysHsAlaSer 281
QY 868 GAGTCGAGAGCCCTGAGTGGCTGAGCAAGAAACCTCCAGAGAGCTGACCTTC 927
DB 282 AspSerAlaIleLeuLeuGlnCysLeuArgAlaValProSerGlnIleLeuMetArgPhe 301
QY 928 AGCCAGAAAACAAGCTTTTCACTCGAGTGTGATGTGCTTTTCTTCTTAATGAGCT 987
DB 302 SerIleValLeuThrPheSerIleProValIleAspAspPhePhePheProAspGlnPro 321
QY 988 CTAGATCTATGTCTCAGAAAGCATTTAAAGCAATTCCTTCAATCTCAGAGTCAATAC 1047
DB 322 ValAlaLeuLeuThrGlnIleValSerAlaPheAsnSerValProSerIleIleIleValAsn 341
QY 1048 CAGAGTGTGCTTCTGCTGCTTGAAGAGGCTCCGAGTCTCTAGTGGCTCAAC 1107
DB 342 HsGlnCysAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlySerAsn 360
QY 1108 AAGTCCCTTGGCTCCATCTGATACAAACATCTGACATCCGCTCAGATTTTGAC 1167
DB 361 ArgSerLeuAlaLeuTyrIleValHsThrPheLeuAsnIleProThrGlnTyrLeuHs 380
QY 1168 CTGTGCTTAATGAACTTCAATGACAAAGCACTCCCTGAATCCGAGACACTTT 1227
DB 381 LeuValAlaAspHsIleTyrPheTyrAsnIleSerProValGlnIleArgAspSerPhe 400
QY 1228 CTGAGCTGTGAGAACTGTCTTGTGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1287
DB 401 LeuAspLeuLeuGlyAspValLeuPheValAlaProGlyValValThrAlaArgTyrHs 420
QY 1288 AGAGATGTGTGTCACCTGTCTACTTCTTGAATTTGGCAGCCGCTCAGTGTGAA 1347
DB 421 ArgAspAlaGlyAlaProValTyrPheTyrGlnPheGlnHsIleProProGlnCysLeuAsn 440
QY 1348 GACACGAGCCGCTTTTGTGTAAGCCGACACGCTGATGAAGTCCGCTTGTGTGCT 1407
DB 441 AspThrArgProAlaPheValIleValSerAlaAspHsIleArgPheValPheGly 460
QY 1408 GTTGTCTTCTGAAGGGGACATTTGTATGTTGGAAGGCGCAGAGAGAGAAATTA 1467
DB 461 GlnAlaPheLeuLeuGlyAspIleValMetPheGlnGlyAlaThrGlnIleGlnGlyLeu 480
QY 1468 CTGAGCCGGAAGATGATGAATTAATCTGAGCTTGTGCTGCAACCGGAACTCTTAATGG 1527
DB 481 LeuSerHsGlyMetMetArgTyrTrpAlaAsnPheAlaArgThrGlyAspProAsnGly 500
QY 1528 AACGACTGTCTGTGCGCACTTAATCTGATCTGACAGTCACTTCAAGTGAAGT 1587
DB 501 GlnGlyValProLeuThrProAlaTyrThrGlnSerGlnGlnIleTyrLeuLeuAspLeu 520
QY 1588 AACATGAGCTTGGAGACAGATCAAGAAACCGCGGTGGAATTTTGGACAGACCATC 1647
DB 521 SerValSerValGlyGlnLeuLeuGlnGlnIleValGlnPheThrMetAsnThrIle 540

RESULT 12
US-10-233-933A-4
Sequence 4, Application US/1023933A
General Information:
Publicat ion No. US20040214171A1
Applicant: Yamashita, Tetsuro
Applicant: Miyazaki, Masao
Title of Invention: CAT KIDNEY DISEASE MARKER
File Reference: SHIG FP02US006
Current Application Number: US/10/233, 933A
Current Filing Date: 2002-09-03
Prior Application Number: JP2002-057908
Prior Filing Date: 2002-04-03
Number of Seq ID Nos: 4
Software: PatentIn Ver. 2.0
Seq ID No 4
Length: 542
Type: PRF
Organism: Felis catus
```

US-10-233-933A-4

Alignment Scores:

Pred. No.:	8,17e-169	Length:	542
Score:	1985.00	Matches:	373
Percent Similarity:	83.27%	Conservative:	60
Best Local Similarity:	71.73%	Mismatches:	85
Query Match:	62.62%	Indels:	2
DB:	16	Gaps:	2

US-10-023-515-3 (1-1746) x US-10-233-933A-4 (1-542)

```
OY 91 GGGCTTGTGTGAAGGGCCACAGAGAAACACAGGCTGGATGATTCAGGGCAAGAA 150
DB 22 GlycerolAcylAAspAlaProValArgSerThrArgLeuGlyTrrValArgGlyysgin 41
OY 151 GTCACTGTGTGGAGAGCCGTGTGCCTGTGAACGTGTTCCTGGAAGTCCCTTGTGCT 210
DB 42 ThrThrValLeuGlySerThrValProValAAspMetPheLeuGly1IleProTyraIaIa 61
OY 211 CCCCCGTGGATCCCTGGATTTAGAAACCCGAGCCGTGCATCCGCCCTGGATAACTTG 270
DB 62 ProProLeuGlyrProLeuArgPheLeuGlnProTyraProValLeuProGlyAAspPhe 81
OY 271 CGAGAGCCACCTCCTACCTAATTGTGTGCTTCAGAACTCAGAGTGGCTGCTTAGAT 330
DB 82 ArgAAspAlaThrSerTyrProTyraLeuGlyrPheGlnAAspLeuGlnTrrPheValSerTyr 101
OY 331 CAACATGTCTCAAGGTGATTAACCCGAAATTCGAGTGTCAAGAACTGCTCTTACCTG 390
DB 102 GlnIleValLeuTyraValArgTyrProTyraLeuGlnAAspGlnAAspTyraLeu 121
OY 391 AACATCTGTGGCCGCGCCAGCCGATACAGGCTCAGAGCTCCGCTGTGGTGTGCTTC 450
DB 122 AAspIleTyraAlaProAlaIleAAspAAspGlySerAAspLeuProValMetValTrrPhe 141
OY 451 CCAGAGGTGCTCTTGAAGACTGCTCAGGCTCATCTTGTGATGGGTCCGCCCTGTGCTGC 510
DB 142 ProGlyGlyAlaPheLeuMetGlySerAlaSerSerPheAAspGlySerAlaLeuAlaIa 161
OY 511 TATGAGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 570
DB 162 TyrrGlnAAspValLeuIleValThrThrGlnTyraGlyLeuGly1IlePheGlyrPheAAsp 181
OY 571 ACATGGGATCAGGATGCTCCGCGGAACTGGGCTTCAAGGACAGAGTGGCTGTGTGCC 630
DB 182 ThrGlyAAspGlnValAlaArgGlyAAspTrrAlaLeuLeuAAspGlnValAlaIaLeuThr 201
OY 631 TGGGTCAAGAAAGATGATGATTTCCGATGGGAGACCCAGCTGTGTGATCACTTTGGGC 690
DB 202 TrrValAAspAAspAAspIleGlnPhePheGlyGlyAAspProAAspSerValThrIlePheGly 221
OY 691 GAGTCCGGGGAGCCATTAAGTGTTCATGCTTAACTGTCTTCCATGGCCAAAGGCTTA 750
DB 222 GlnSerAlaGlyAlaIleSerValSerSerLeuIleLeuSerProIleAlaAAspGlyLeu 241
OY 751 TTCCCAAGGATGATGAGAGAGTGGGAGGCTCATCTCCCTTACCTG---GAGGCCAT 807
DB 242 PheHisValValAlaIleMetGlnSerGlyValAlaIleLeuProLeuLeuMetAAspProPro 261
OY 808 GATTATGAGAAAGTGAAGACTGAGAGTGTGTGACATTTCTGTGTAAACATGCTCA 867
DB 262 GlyAAspGlnAAspGlyrPheLeuGlnValLeuAlaAAspGlyCyGlyGlyHisAlaAAsp 281
OY 868 GACTTGAAGGCTCTGTGAGTGTGTGAGACAAACCTTCAAGAGAGTGTGTGACCTTC 927
DB 282 AAspSerAlaAlaLeuLeuGlnCysLeuAAspGlyAAspProSerGlnGlnLeuMetAAspIle 301
OY 928 AGCGAAGAAACAAATCTTTCACGTGAGTGTGTGATGGCTTCTTCTTCAATGAGCT 987
DB 302 SerTyraValLeuThrPheSerIleProValIleAAspAAspPhePhePheProAAspGlnPro 321
OY 988 CTAGATCTATTGTCTCAGAAAGCATTTAAAGCAATTCTTCATGATCGAGTCAATAC 1047
```

```
DB 322 ValAlaLeuLeuThrGlnIleValAlaPheAAspSerValProSerIleIleGlyValAAspAAsp 341
OY 1048 CACGAGTGTGCTTCTGTGCTGCTTATGAGAGGCTCTGTGAGTCTTCAAGTGTCCCAAC 1107
DB 342 HisGlnCysAlaPheLeuLeu---SerThrGlnPheSerGlnIleLeuGlyGlySerAAsp 360
OY 1108 AAGTCCCTGGCCCTTCATCTGATACAAACATCCGAGCATCCGAGCTCAGATTTGAC 1167
DB 361 ArgSerLeuAlaLeuTyraLeuValHisThrPheLeuAAspIleProThrGlnTyraIleHis 380
OY 1168 CTGTGTGCTTAATGATTAATCTTCATGACAAAGCACTCCCTGACTGAATTCGAGACAGTCTT 1227
DB 391 LeuValAlaAAspHisTyraPheTyraAAspHisSerProValGlnIleAAspSerPhe 400
OY 1228 CTGACACTGTGTGAGATGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1287
DB 401 LeuAAspLeuLeuGlyAAspValLeuPheValAlaProGlyValAlaThrAlaArgTyraHis 420
OY 1288 AGAGATGTGTGGACCTGTCTACTTCTPATGAGTTTCGAGACCGGCTCAGTGTCTTGA 1347
DB 421 ArgAAspAlaGlyAlaProValTyrPheTyrrGlnPheGlnHisProProGlnCysLeuAAsp 440
OY 1348 GACACGAAGCCGCTTTTGTCAAAAGCCAGCAAGCTGATGAAGTCCGCTTGTGTGCT 1407
DB 441 AAspThrArgProAlaPheValIleAAspHisSerAAspGlnIleArgPheValPheGly 460
OY 1408 GTGTGCTTCTGGAAGGGGACATTTGATTTCCAGAAAGCCAGGAGAGAGAAAGTTA 1467
DB 461 GlyAlaPheLeuTyraGlyAAspIleValMetPheGlnGlyAlaThrGlnGlnGlyLeu 480
OY 1468 CTGAGCCGGAAGATGATTAATCTGAGCTACTTGTGTGCAACCGGGAATCTTATGGG 1527
DB 481 LeuSerArgGlyMetMetAAspTyrTrrAlaAAspPheAlaAAspThrGlyAAspProAAspGly 500
OY 1528 AACGACTGTCTGTGTGCGACGTTTAAATCTGACTGAGACATCTTCAGCTGACTTG 1587
DB 501 GlnGlyValProLeuTrrProAlaTyrThrGlnSerGlnGlnTyrrLeuTyraLeuAAspLeu 520
OY 1588 AACATGAGCTCCGAGACAGAGACTCAAAAGCCGCGGTGAGATTTTGGACAGACCACTC 1647
DB 521 SerValSerValGlyGlnTyraLeuGlyGlnGlnGlnValGlnPheTyraMetAAspThrIle 540

RESULT 13
US-10-451-168-93
; Sequence 93, Application US/10451168
; Publication No. US20040091969A1
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; APPLICANT: GLAXO GROUP LIMITED
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FIDR REFERENCE: GFS0039
; CURRENT APPLICATION NUMBER: US/10/451,168
; CURRENT FILING DATE: 2003-11-12
; PRIOR APPLICATION NUMBER: PCT/US01/49232
; PRIOR FILING DATE: 2000-12-17
; PRIOR APPLICATION NUMBER: 60/256,710
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/257,048
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/260,482
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/264,922
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/266,797
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/276,988
; PRIOR FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/281,535
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/289,622
; PRIOR FILING DATE: 2002-06-28
```

```

; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 93
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-451-168-93

Alignment Scores:
Pred. No.: 1 74e-150 Length: 356
Score: 1780.00 Matches: 340
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 56.15% Indels: 0
DB: 15 Gaps: 0

US-10-023-515-3 (1-1746) x US-10-451-168-93 (1-356)

QY 724 ATACTGCTCCAGGCGCAAGGCTTATTCACAAAGCCATGATGAGAGTGCGGCGC 783
DB 17 ILeuSerProMetAlaIysGlyLeuPheHisIlySalalIeMetGluSerGlyValAla 36
QY 784 ATCATCTCCCTTACCTGAGGCCCAATGATTATGAGAGAGTGAGGAGCTTGAGTGCTGCA 843
DB 37 IleIleProIyLeuGluIahIHisAspIyGlyIySerGluAspLeuGlnValAla 56
QY 844 CATTTCTGTGTTAAACAATCCGTCAAGCTCTGAGGCCCTGCTGAGGCTTGAGGAGAAA 903
DB 57 HisPheCyGlyIAsnAsnAlaSerAspSerGluAlaLeuIleuArgCyLeuIyGlnIy 76
QY 904 CCCTCAAGAGAGCTGCTGAGCCCTCAAGCCAGCAAAAACAAGCTTTCTAGCTGAGTGAT 963
DB 77 ProSerIyGlyLeuLeuThrLeuSerGlnIyThrIySerPheThrArgValAlaAsp 96
QY 964 GGTGCTTTCTTCTTATGAGCCCTTATGATCTATTGTCTCAGAAAGCAATTAAAGCAATT 1023
DB 97 GlyAlaPhePheProAsnGluProLeuAspLeuLeuSerGlnIySalalIe 116
QY 1024 CCTTCATATATGAGATCAATAACCAAGATGGCTTCTGCTGCTTATGAAAGAGGCT 1083
DB 117 ProSerIleIleGlyValAsnAsnHisGlyIyGlyPheLeuLeuProMetIyGlyAla 136
QY 1084 CCTGAGATCCTGAGGCTCAACAAGGCTCCCTGCTCATCTGATACAAACATCTG 1143
DB 137 ProGluIleuSerGlySerAsnIySerLeuAlaLeuHisIleuIleGlnAsnIleu 156
QY 1144 CACATCCCGCTCAGATTATGCACTTGTGGCTTATGAAATCTTCCATGACAGCACTCC 1203
DB 157 HisIleProProGlnIyLeuHisIleuValAlaAsnGluIyPheHisAspIyHisSer 176
QY 1204 CTGACTGAAATCCGAGACAGTCTTCTGACTGCTTGGAGATGTGTTTGTGGTCCCT 1263
DB 177 LeuThrGluIleArgAspSerLeuLeuAspLeuLeuGlyAspValPhePheValAlaPro 196
QY 1264 GCACTGATCAGCTCGATATACAGAGATGCGTGGCACTGCTTCACTTATGAGTTT 1323
DB 197 AlaLeuIleThrAlaArgIyTrHisAspAspAlaGlyAlaProValIyTrPheIyGlyIyPhe 216
QY 1324 CGGACCGGCTCAGTCTTGTGAAGACAGAAAGCCGGCTTTGTCAAGCCGACAGCT 1383
DB 217 ArgHisArgProGlnIySerPheGluAspThrIyIyProAlaPheValIySalalIeAspHis 236
QY 1384 GATGAAGTCCGCTTGTGTGGTGGTGGCTTCTGAAGGGGAGCACTGTATGTTGGA 1443
DB 237 AspGluValAlaArgPheValPheGlyGlyAlaPheLeuIyGlyAspIleValaMetPheGlu 256
QY 1444 GGAAGCAGGAGAGAGAGAGATTACTGAGCCGGAAGATGATGAANAATCGAGTACTTT 1503
DB 257 GlyAlaThrGluGluGluIyIyLeuLeuSerArgIyIyMetCylSerIyTrIyAlaThrPhe 276
QY 1504 GCTCGAACCAGGAATCTTATGAGGAACGACTGTCTGTGGCCAGCTTATATCTGACT 1563
DB 277 AlaArgThrGlyIyAsnProAsnGlyIyAsnAspLeuSerLeuIyTrIyProAlaIyTrIyAsnLeuThr 296

```

```

QY 1564 GAGCAGTACTTCAGCTGATCTTGAACATGAGCCTTCGACAGAGACTCAAGAACCGCGG 1623
DB 297 GluGlnIyLeuGlnIleuAspLeuAsnMetSerLeuGlyGlnArgLeuIyGluProArg 316
QY 1624 GTGATTTTGGACGACGACCATCCCTGATCTGTCTGCTCGACATGCTCCCACT 1683
DB 317 ValAspPheTrpThrSerThrIleProLeuIleuSerAlaSerAspMetLeuHisSer 336
QY 1684 CCTCTTCTCTTAACTTCTCTCTCTCTCCAGCCCTTCTTTCTTTGTGCTCCT 1743
DB 337 ProLeuSerSerLeuThrPheLeuSerLeuGlnIyProPhePhePheCysAlaPro 356

RESULT 14
US-09-925-298-689
; Sequence 689, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4103
; CURRENT APPLICATION NUMBER: US/09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 689
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-689

Alignment Scores:
Pred. No.: 5 09e-101 Length: 549
Score: 1230.00 Matches: 248
Percent Similarity: 63.84% Conservative: 91
Best Local Similarity: 46.70% Mismatches: 168
Query Match: 38.80% Indels: 24
DB: 9 Gaps: 8

US-10-023-515-3 (1-1746) x US-09-925-298-689 (1-549)

QY 91 GGGCTTCTGTAAGGCGCACAGAGAACACAGGCTGGATGATTCAGGGCAAGCA 150
DB 16 GlyGlnAspSerAlaSerProIleArgThrTrHisIleThrGlyGlnValLeuGlySerLeu 35
QY 151 GTCACTGTCTGGGAAGCCCTGCTGTGAACAGTCTCTGGAGTCCCTTGTGCT 210
DB 36 ValHisValIyGlyAlaAsnAlaGlyValGlnThrPheLeuGlyIleProPheAlaIy 55
QY 211 CCCCCCTGGATCCCTGCGATTTACGAACCCGACGCTGCATCGCCCTGGGATACTTG 270
DB 56 ProProLeuGlyIyProLeuArgPheAlaProProGluIyProGluSerIyPheSerGlyVal 75
QY 271 CGAAGACCACTCTTACCTTAATTGTGCTCTCAGAAC-----TCAGAG 315
DB 76 ArgAspGlyThrThrHisProAlaMetCylLeuGlnAspLeuThrAlaValGluSerGlu 95
QY 316 TGGCTGCTTATGATCAACAACATGCTCAAGGTGCATTACCGAAATTCGAGTTCAGAA 375
DB 96 PheLeu-----SerGlnPheAsnMetThrPheProSerAspSerMetSerGln 111
QY 376 GACTGCTTACTGAAACATCTATGCGCTGCGCCAGCGCATTAAGCTCCCAAGCTCCCC 435

```



```

Db      96 PheLeu-----SerGlnPheAsnMetThrPheProSerAspSerMetSerGln 111
Qy      376 GAGTGGCTCTACCTGAACATCTATGCGCTGGCCACGCCGATACAGGCTCCAAAGCTCCCC 435
Db      112 AspCysLeuTyrLeuSerIleTyrThrProAlaHisSerHisGlnGlySerAsnLeuPro 131
Qy      436 GTCTTGGTGTGTGTCCCAAGAGGTGCTTCAGAGCTGCGCTCAGCTTCATCTTTGAATGGG 495
Db      132 ValMetValTPrIleHisGlyGlyAlaLeuValPheGlyMetAlaSerLeuTyrAspGly 151
Qy      496 TCCGCGCTGGCTGCTATGAGAGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
Db      152 SerMetLeuAlaAlaLeuGlnAsnValValValIleIleGlnTyrArgLeuGlyVal 171
Qy      556 TTTGGTCTTCACCAACATGGGATCAGCATGCTCGGGGAACTGGAGCCCTCAAGGACGAG 615
Db      172 LeuGlyPhePheSerThrGlyAspTyrHisAlaThrGlyAsnTyrGlyTyrLeuAspGln 191
Qy      616 GTGGCTGCTGTCTGCTGGGTCAGAGAAACATCGAGTTCTTGCTGGTGGGACCCAGCTCT 675
Db      192 ValAlaAlaLeuArgTPrValGlnGlnAsnIleAlaHisPheGlyGlyAsnProAspArg 211
Qy      676 GTGACCATCTTTGGCGAAGTCGCGGAGCCATAGTGTCTTCTTATCTGTCTGCC 735
Db      212 ValThrIlePheGlyGlySerAlaGlyTyrThrSerValSerSerLeuValValSerPro 231
Qy      736 ATGGCCAAAGCTTATTCACAAAGCCATGAGAGAGTGGGGTGGCCATCATCCCTTAC 795
Db      232 IleSerGlnGlyLeuPheHisGlyAlaIleMetGlnSerGlyValAlaLeuLeuProGly 251
Qy      796 CTGAGAGCCCATGATTATGAGAAAGTGAGACTG-----CAGGTGGTTGCACATTTTC 849
Db      252 LeuIleAla-----SerSerAlaAspValIleSerThrValValAlaAsnLeu 267
Qy      850 TGTGTGAACATGCGCTCAGACTCTGAGGCTGTGAGGTCCTGAGGACAAACCTTCC 909
Db      268 SerAlaCysAspGlnValAlaAspSerGlnAlaLeuValGlyCysLeuArgGlyLysSerLys 287
Qy      910 AAGAGCTGTCTGACCTCAGCCAGCAAGAAACAAAGCTTTCACTCGAGTGTGTGTGTGTGT 969
Db      288 GlnGlnIleLeuAlaIleAsnLysProPheLysMetIleProGlyValValAspGlyVal 307
Qy      970 TTCTTTCCATGAGCCTCTAGATCTATGTCCTCAGAAAGCATTTAAAGCAATTCCTTCC 1029
Db      308 PheLeuProAlaGlnProGlnGlnLeuLeuAlaSerAlaAspPheGlnProValProSer 327
Qy      1030 ATCATCGAGATCAATTAACACAGATGTGTGCTTCTGCTGCT-----ATGAG 1077
Db      328 IleValGlyValAlaAsnAsnAsnGlnPheGlyTPrLeuIleProLysValMetArgIleTyr 347
Qy      1078 GAGGCTCTGAGATCTCTCAGTGGCTCCAAAGATCCCTTGCTTCATCTGATACAAAC 1137
Db      348 AspThrGlnLysGlnMetCAspArgGlnAlaSerGlnAlaAlaLeuGlnLysMetLeuThr 367
Qy      1138 ATCTCGACATCCGCGCTCAGATTTTGCACTTTGGGCTAATGAAATCTTCATGACAG 1197
Db      368 LeuLeuMetLeuProThrPheGlyAspLeuLeuArgGlnGlnTyrIleGlyAspAsn 387
Qy      1198 CACTCCCTGACTGAATCCGAGACAGTCTTCTGACTTGCTTGAGATGTGTTCTTTGTG 1257
Db      388 GlyAspProGlnThrLeuGlnAlaGlnPheGlnGlnGlnMetCAlaAspSerMetPheVal 407
Qy      1258 GTCCCTGACATGATCAACAGCTCGATATTCACAGAGATGCTGTGACCTGTCTTACTTAT 1317
Db      408 IleProAlaLeuGlnValAlaHisPhe---GlnCysSerArgAlaProValTyrPheTyr 426
Qy      1318 GAGTTTGGCAACGGGCTCAGAGCTTTGAAGACAGAAAGCGGCTTTTGTCAAAGCCGAC 1377
Db      427 GlnPheGlnHisGlnProSerTPrLeuLysAsnIleArgProProHisMetLysAlaAsp 446
Qy      1378 CAGCGTGAATGAAGTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATG 1437
Db      447 HisGlyAspGlnLeuProPheValPhe---ArgSerPhePheGlyGlyLysAsnTyrIleLys 465

```

```

Qy      1438 TTGGAAGAGCCACGAGAGAGAGAAATTATCTGACCCGAGAGATGATAAATATCTGGCT 1497
Db      466 Phe-----ThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 482
Qy      1498 ACCTTTGTGCAACCGGAAATCTTAATGGGAAAGACCTGTCTGTGTGTGTGTGTGTGTGT 1557
Db      483 AsnPheAlaArgAsnGlyAsnProAsnGlyGlnGlyLeuProHisIleTPrProLeuPheAsp 502
Qy      1558 CTGACTGAGCGATACCTCCAGCTGAGCTTGAACATGAGCCTGGACAGAGACTCAAGAA 1617
Db      503 GlnGlnGlnGlnTyrLeuGlnLeuAsnLeuGlnProAlaValGlyArgAlaLeuLysAla 522
Qy      1618 CCGCGGTGATTTTGTGACCAAGACCAATCCCC 1650
Db      523 HisArgLeuGlnPheTPrLysAlaLeuPro 533

```

Search completed: June 15, 2005, 09:37:11
 Job time : 152.806 secs

This Page Blank (uspto)